

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: August 4, 2003, 21:01:54 ; Search time 7372.88 Seconds

(without alignments)  
11491.278 Million cell updates/sec

Title: US-09-936-271B-13\_COPY\_9500\_11570

Perfect score: 2071

Sequence: 1 aaacagaccacaactctct.....ccacgaataactggaag 2071

Scoring table: IDENTITY\_NNC

Gapop 10.0 , Gapept 1.0

Searched: 2888711 segs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_iny:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pla:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_man:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2071	100.0	11570	9	AF135028	AF135028 Homo sapi
2	2071	100.0	107487	9	AC011483	AC011483 Homo sapi
3	2071	100.0	230000	9	AF243527	AF243527 Homo sapi
4	1990	96.1	217346	2	AC027602	AC027602 Homo sapi
5	1686.8	81.4	200792	2	AC130782	AC130782 Pan trogl
6	482	23.3	586	6	AX331407	AX331407 Sequence
7	482	23.3	586	6	AX331824	AX331824 Sequence
8	482	23.3	586	11	G41906	G41906 SHG-56840
9	482	23.3	735	6	AX429955	AX429955 Sequence
10	482	23.3	738	6	BD139877	BD139877 Compounds
11	482	23.3	1370	9	AY279380	AY279380 Homo sapi
12	482	23.3	1381	6	BD107879	BD107879 36 human
13	482	23.3	1387	9	AF168768	AF168768 Homo sapi
14	482	23.3	1438	9	AY279381	AY279381 Homo sapi
15	482	23.3	1439	6	BD107880	BD107880 36 human
16	482	23.3	1499	6	BD005362	BD005362 Protease.
17	482	23.3	1516	6	BD107865	BD107865 36 human
18	482	23.3	1536	9	BC008036	BC008036 Homo sapi
19	482	23.3	1570	6	AR252580	AR252580 Sequence
20	482	23.3	1570	6	AX080829	AX080829 Sequence
21	482	23.3	1570	6	AX403421	AX403421 Sequence
22	482	23.3	1570	6	AX464322	AX464322 Sequence
23	476	23.0	1476	6	AR078184	AR078184 Sequence
24	476	23.0	1476	6	AR137506	AR137506 Sequence
25	476	23.0	1476	6	AR242358	AR242358 Sequence
26	476	23.0	1476	6	BD082659	BD082659 Keratinoc
27	460	22.2	1504	6	AR263823	AR263823 Sequence
28	449	21.7	673	6	AX067353	AX067353 Sequence
29	270.4	13.1	57729	9	AL449284	AL449284 Human DNA
30	264.2	12.8	42107	9	AC004409	AC004409 Homo sapi
31	261.2	12.6	129048	9	AC020931	AC020931 Homo sapi
32	260	12.6	120997	9	AC013449	AC013449 Homo sapi
33	255.8	12.4	23615	2	AP000933	AP000933 Homo sapi
34	252.2	12.2	118085	9	AL136115	AL136115 Human DNA
35	250.8	12.1	172945	9	AC007220	AC007220 Homo sapi
36	249.8	12.1	160915	2	AP001084	AP001084 Homo sapi
37	249.8	12.1	165420	9	AC091588	AC091588 Homo sapi
38	249.8	12.1	171978	2	AC009669	AC009669 Homo sapi
39	249.8	12.1	204504	2	AC022487	AC022487 Homo sapi
40	249.8	12.0	208310	9	AC091043	AC091043 Homo sapi
41	249	11.9	190982	2	AP002344	AP002344 Homo sapi
42	247.4	11.9	190982	2	AC025285	AC025285 Homo sapi
43	246.8	11.9	160858	9	AC019216	AC019216 Homo sapi
44	246	11.9	152659	9	AL591503	AL591503 Human DNA
45	243.2	11.7	40395	9	AC099491	AC099491 Homo sapi

## ALIGNMENTS

RESULT 1  
AF135028  
LOCUS AF135028  
DEFINITION Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.  
ACCESSION AF135028  
VERSION AF135028.1 GI:4589282  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 11570)  
TITLE The new kallikrein-like gene, KLK-L2. Molecular characterization,  
mapping, tissue expression, and hormonal regulation

JOURNAL MEDLINE  
J. Biol. Chem. 274 (53), 37511-37516 (1999)  
10608802  
2 (bases 1 to 11570)  
Diamandis, E.P., Yousef, G.M., Luo, L.Y., Magklara, A. and Oblezu, C.V.  
The new human kallikrein gene family: implications in  
carcinogenesis  
Trends Endocrinol. Metab. 11 (2), 54-60 (2000)  
JOURNAL MEDLINE  
21121728  
10675891  
3 (bases 1 to 11570)  
Yousef, G.M., Luo, L.Y. and Diamandis, E.P.  
Direct Submission  
Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount  
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,  
Canada

FEATURES  
source Location/Qualifiers  
1..11570  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.3-q13.4"  
join(2101..2131,2210..2293,4762..5023,5763..6019,  
6105..6238,11092..11570)  
/product="kallikrein-like protein 2 KLK-L2"  
join(2221..2293,4762..5023,5763..6019,6105..6238,  
11092..11247)  
/codon\_start=1  
/product="kallikrein-like protein 2 KLK-L2"  
/protein\_id="AAD26429.1"  
/db\_xref="GI:4589283"  
/translation="MATAAPPMMWVLCALITALLGVTEHYLVANDVSCDHPNTPVS  
/GENOLIGAGGEDARDSDSSRLINGSDDMDHTOPMOALIRPOLYCGALVHPOM  
LITPAHCKRVKRVRLGHYSIPVYESGOMQGVKSIIPHPYSHRHSNLMILKIN  
RIIRPTKDVRLPYNVSSHCPSSAKTCLVSWMTKSPQVHFVLOCINISVLSQKRC  
DAYPRQIDDTMFCADKAGRDSCQDSDSGPVVNGSLQGLVSWGDPICAPRNRPVYT  
NLCKETKWOEIFIQANS"

BASE COUNT 3034 a 2562 c 3325 g 2649 t  
ORIGIN

Query Match 100.0%; Score 2071; DB 9; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2071; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACAGACCAAAACCTTCCTCCGCGCTGAGACCTCATGTGCCCAAGTGAAGACAGCAA 60  
DB 9500 AAACAGACCAAAACCTTCCTCCGCGCTGAGACCTCATGTGCCCAAGTGAAGACAGCAA 9559  
QY 61 TAAAGATAGATTAATATGTAGTAATTAATAAAAAAAAAAAAAAAAAATTAAGCCGGGTGTGGT 120  
DB 9560 TAAAGATAGATTAATATGTAGTAATTAATAAAAAAAAAAAAAAAAAATTAAGCCGGGTGTGGT 9619  
QY 121 CTTCGACCTGTAGTTCACACTACTTGGAGGCTGAGAGTGGAGAAATTCGTGAGCCCAA 180  
DB 9620 CTTCGACCTGTAGTTCACACTACTTGGAGGCTGAGAGTGGAGAAATTCGTGAGCCCAA 9679  
QY 181 CGTTTAGAGCTGGGTAAGCACTGACTGCTGCTCAGTCCAGACAGCAGCTGGGTGA 240  
DB 9680 CGTTTAGAGCTGGGTAAGCACTGACTGCTGCTCAGTCCAGACAGCAGCTGGGTGA 9739  
QY 241 CAAAGCAAGACGTTTTTGTCAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
DB 9740 CAAAGCAAGACGTTTTTGTCAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9799  
QY 301 AAAG 360  
DB 9800 AAAG 9859  
QY 361 GAAAG 420  
DB 9860 GAAAG 9919

QY 421 GAT 480  
DB 9920 GAT 9979  
QY 481 TATTATAGAGAGGTAATATATAGAGAGGTAATAGAGAGGTAATAGAGAGGTAATAGAGAG 540  
DB 9980 TATTATAGAGAGGTAATATATAGAGAGGTAATAGAGAGGTAATAGAGAGGTAATAGAGAG 10039  
QY 541 TCAGAGCAATATGATTTCTATTTGGAGTATTTGCCCCCTAGAGAGACACTGGCAATACCA 600  
DB 10040 TCAGAGCAATATGATTTCTATTTGGAGTATTTGCCCCCTAGAGAGACACTGGCAATACCA 10099  
QY 601 GGAGACATTTTGGTTGTCACACATATATGAGAGGCACTTCTGCAACTATGATATGA 660  
DB 10100 GGAGACATTTTGGTTGTCACACATATATGAGAGGCACTTCTGCAACTATGATATGA 10159  
QY 661 TGCCCAAGTGTGCTGTCAACATGCTATGATGACACGGCCAGGCTCCCAACCAACCTT 720  
DB 10160 TGCCCAAGTGTGCTGTCAACATGCTATGATGACACGGCCAGGCTCCCAACCAACCTT 10219  
QY 721 ATCCAGCTTCAGATGCCAGATGCCAGATCCAGAGAACCCCTCATCCAGGGGCTGAGAAC 780  
DB 10220 ATCCAGCTTCAGATGCCAGATGCCAGATCCAGAGAACCCCTCATCCAGGGGCTGAGAAC 10279  
QY 781 CGTATTTTTCAGAGAGAGAGATATAGAGATGGGTTGGTGAGAGATGGGAGAGAGAGATGT 840  
DB 10280 CGTATTTTTCAGAGAGAGAGATATAGAGATGGGTTGGTGAGAGATGGGAGAGAGAGATGT 10339  
QY 841 GTGTCCAGTAAGAGAAATATAGCCCTGACAGAGCTGAGAGGGAGATGAGAGAGAGAGAG 900  
DB 10340 GTGTCCAGTAAGAGAAATATAGCCCTGACAGAGCTGAGAGGGAGATGAGAGAGAGAGAG 10399  
QY 901 GGCAGAGAGATACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 960  
DB 10400 GGCAGAGAGATACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 10459  
QY 961 GATTTGGAGAGAGAGAGGTCACAGAGACCCCGCAAAATGATGTGTGACAAACAGATCTGG 1020  
DB 10460 GATTTGGAGAGAGAGAGGTCACAGAGACCCCGCAAAATGATGTGTGACAAACAGATCTGG 10519  
QY 1021 AAGAGAGATGAGTGGAGAGTGAACAATGGGGCTCTAAGGTTGAACCTTGGAGGCCAGG 1080  
DB 10520 AAGAGAGATGAGTGGAGAGTGAACAATGGGGCTCTAAGGTTGAACCTTGGAGGCCAGG 10579  
QY 1081 CATGGTGGCTCAGCGCTGTATTCACACACTTTGAGAGCTGAGTGGGCGAATCACTTGA 1140  
DB 10580 CATGGTGGCTCAGCGCTGTATTCACACACTTTGAGAGCTGAGTGGGCGAATCACTTGA 10639  
QY 1141 GGCAGAGAGTTCGAGACCAAGCCTGGCCAAATGGTGAACCCCGTCTACAAAAAAAT 1200  
DB 10640 GGCAGAGAGTTCGAGACCAAGCCTGGCCAAATGGTGAACCCCGTCTACAAAAAAAT 10659  
QY 1201 ACAAATAATTTAGCCGGCTGTGTGATGACACTGTAGTACAGCTACTTGGAGGCTGA 1260  
DB 10700 ACAAATAATTTAGCCGGCTGTGTGATGACACTGTAGTACAGCTACTTGGAGGCTGA 10759  
QY 1261 GGCAGAGAGATTCGTTGAACCCGAGAGATGAGAGCTGACGTGAGCTGAGCTCAGCCACT 1320  
DB 10760 GGCAGAGAGATTCGTTGAACCCGAGAGATGAGAGCTGACGTGAGCTGAGCTCAGCCACT 10819  
QY 1321 GCGCTCAACCTGGGCAACAGAGTAAGACTCCATCTCAAAAAAAGAGAGAGAGAGAGAG 1380  
DB 10820 GCGCTCAACCTGGGCAACAGAGTAAGACTCCATCTCAAAAAAAGAGAGAGAGAGAGAG 10879  
QY 1381 AGTGAATATTTATATACATTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440  
DB 10880 AGTGAATATTTATATACATTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10939  
QY 1441 TTTTTCGATTTCTTCATCTCTGATCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500  
DB 10940 TTTTTCGATTTCTTCATCTCTGATCTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10999  
QY 1501 TCCATCTATGGGATCTCTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560



```

Db 49940 CGTATTTTTCAGAGAGGAGGTATTAAGATGGGTTGTGTGAGAAATGGGAAAGAGGTGT 49881
Oy 841 GTGTCCATTAAGAGAAATTAAGCCCTGCACAGGCTGGAGGGAGATGAGAGAAAGGA 900
Db 49880 GTGTCCATTAAGAGAAATTAAGCCCTGCACAGGCTGGAGGGAGATGAGAGAAAGGA 49821
Oy 901 GGGGAGAGATACAGATGAGGGAGACAGGCTGGAACAGAAAGTAGACAGCAAGATTGA 960
Db 49820 GGGGAGAGATACAGATGAGGGAGACAGGCTGGAACAGAAAGTAGACAGCAAGATTGA 49761
Oy 961 GATGTGGAGAGAGAGGTACACAGACCCCGGAATGATGTGTGACAAACAGATCTGG 1020
Db 49760 GATGTGGAGAGAGAGGTACACAGACCCCGGAATGATGTGTGACAAACAGATCTGG 49701
Oy 1021 AAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1080
Db 49700 AAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 49641
Oy 1081 CATGTGGCTCAGCGCTGTAAATCCCAACACTTTGGAGGCTGAGTGGCGAATCACTGA 1140
Db 49640 CATGTGGCTCAGCGCTGTAAATCCCAACACTTTGGAGGCTGAGTGGCGAATCACTGA 49581
Oy 1141 GGCCAGAGTTTCAGAGACAGGCTGGCCCAACATGCTGAACCCCTCTCTACAAAAAAAT 1200
Db 49580 GGCCAGAGTTTCAGAGACAGGCTGGCCCAACATGCTGAACCCCTCTCTACAAAAAAAT 49521
Oy 1201 ACAAATAATTAGCCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 49520 ACAAATAATTAGCCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 49461
Oy 1261 GGCAGAGAAATTTGCTTGAACCCCGGAGATGAGAGCTGCAGTGTGAGTGTGAGCTCACT 1320
Db 49460 GGCAGAGAAATTTGCTTGAACCCCGGAGATGAGAGCTGCAGTGTGAGTGTGAGCTCACT 49401
Oy 1321 GCGCTCCAACTGGGCAACAGAGTAAGTCCATCTCAAAAAAAAGCTGGATTGG 1380
Db 49400 GCGCTCCAACTGGGCAACAGAGTAAGTCCATCTCAAAAAAAAGCTGGATTGG 49341
Oy 1381 AGTGAATATTAATTAACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
Db 49340 AGTGAATATTAATTAACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 49281
Oy 1441 TTTTGTGATTTCTTCACTCTGTACTTTCATCTCTGTGTGTGTGTGTGTGTGTGTGTGT 1500
Db 49280 TTTTGTGATTTCTTCACTCTGTACTTTCATCTCTGTGTGTGTGTGTGTGTGTGTGTGT 49221
Oy 1501 TCCATCTAATGGGCACTCTGGGCTCTCATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560
Db 49220 TCCATCTAATGGGCACTCTGGGCTCTCATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 49161
Oy 1561 GCGCTCTCATGCCCCCTTTCTCTCTCTGACAGGATGATTTGGGGGCTGTGTGTGTGA 1620
Db 49160 GCGCTCTCATGCCCCCTTTCTCTCTCTGACAGGATGATTTGGGGGCTGTGTGTGTGA 49101
Oy 1621 ATGCTCTCCCTGACAGGACTCTGTGTCTCTGGGAGATTACCTTGTGCGGCGCCAAACAG 1680
Db 49100 ATGCTCTCCCTGACAGGACTCTGTGTCTCTGGGAGATTACCTTGTGCGGCGCCAAACAG 49041
Oy 1681 CGGCTGTCTACAGAACTCTGCAAGTTTACCAAGTGTGTCAGGAAACATCCAGGGCA 1740
Db 49040 CGGCTGTCTACAGAACTCTGCAAGTTTACCAAGTGTGTCAGGAAACATCCAGGGCA 48981
Oy 1741 ACTCTGAGTATCTCCAGAGACTCAGACACCGGCACTCCCACTCTGAGGAGACGCC 1800
Db 48980 ACTCTGAGTATCTCCAGAGACTCAGACACCGGCACTCCCACTCTGAGGAGACGCC 48921
Oy 1801 TGACACTCTCTTCAAGACCTCAATCTCTTCCAGAGATGTTGAATGTTTCACTCTCCAG 1860
Db 48920 TGACACTCTCTTCAAGACCTCAATCTCTTCCAGAGATGTTGAATGTTTCACTCTCCAG 48861
Oy 1861 CCCCTGACCCCATGTCTCTGAGTCAAGGTTCTGCTCCCACTTGGGCTGACCGTGT 1920
Db 48860 CCCCTGACCCCATGTCTCTGAGTCAAGGTTCTGCTCCCACTTGGGCTGACCGTGT 48801

```

```

Oy 1921 CTCTCTAGTTGACCCCTGGGAACAATTTCCAAAACCTGTCAGGGCGGGGTTGGCTCTCA 1980
Db 48800 CTCTCTAGTTGACCCCTGGGAACAATTTCCAAAACCTGTCAGGGCGGGGTTGGCTCTCA 48741
Oy 1981 ATCTCCCTGGGGCACTTTCATCTCAAGCTTCAAGGCGCCATCTCTTCTGTGAGCTGTGAC 2040
Db 48740 ATCTCCCTGGGGCACTTTCATCTCAAGCTTCAAGGCGCCATCTCTTCTGTGAGCTGTGAC 48681
Oy 2041 CCAATTATGTCGCAAAATTAACAGAGAG 2071
Db 48680 CCAATTATGTCGCAAAATTAACAGAGAG 48650

RESULT 3
LOCUS AF243527/c 230000 bp DNA linear PRI 21-NOV-2000
DEFINITION Homo sapiens serine protease gene cluster, complete sequence.
ACCESSION AF243527
VERSION AF243527.1 GI:11244757
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE Gan,L., Lee,I., Smith,R., Argonza-Barrett,R., Lei,H., McCuaig,J., Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS Gan,L., Lee,I., Smith,R., Argonza-Barrett,R., Lei,H., McCuaig,J., Moss,P., Paepert,B. and Wang,K.
TITLE Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region
JOURNAL Gene 257 (1), 119-130 (2000)
MEDLINE 20510030
PUBMED 11054574
REFERENCE 2 (bases 1 to 230000)
AUTHORS Gan,L., Lee,I., Smith,R., Argonza-Barrett,R., Lei,H., McCuaig,J., Moss,P., Paepert,B. and Wang,K.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE, Bothell, WA 98021, USA
FEATURES
source
1..230000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
/gene="KLK1"
/complement(join(<6963..>11517))
/gene="KLK1"
/complement(join(<6963..7118,7668..7804,7923..8212,9482..9641,11472..>11517))
/gene="KLK1"
/product="renal kallikrein"
/complement(join(6963..7118,7668..7804,7923..8212,9482..9641,11472..11517))
/gene="KLK1"
/note="serine protease"
/codon_start=1
/product="renal kallikrein"
/protein_id="AAG33353.1"
/db_xref="GI:11244758"
/translation="MWPLVLCGLASLSTGTAAPPIQRIYGVHCEQHSOPWMAATYH FSTPQGGILVHROWLVLAHICISDNTQVLMVGRNLTPDENTAQFVHSEFPIRGN MLELNHTROADEDVSHDLMLRLTEPADTITDAVKVLEPTEPEVSGTSLASGWS IEPEVSEPDLOCVDLKLILPNDECKKHHVQKVDPMVLGVHLEGGKDTGVDSGGL MCDVIGVTSNGVPGSTPNKPSVAIVKLSYVWIEDTILENS"
/complement(join(<13552..13704,14377..14530,14678..14917,15416..15569,19204..>19246))
/product="ACO protease"
/complement(join(13552..13704,14377..14530,14678..14917,15416..15569,19204..>19246))
/note="serine protease"
/codon_start=1
/product="ACO protease"

```



/protein\_id="AAG33354.1"  
 /db\_xref="GI:11244759"  
 /translation="MMLLTLSLESLASTAADDKLEGBEACPHSPQMOVALYERGR  
 FNCASLISHWVLSAAHCQSRERKVLGHNLRKRDGPDLRTTSRVITPHPREAAS  
 HRNDIMLRVOPARLNPQVPAVLPTRCPHGPACVAGSGMLPSSPYSLPPTLIC  
 ANSIISDSCDKSYPGRLTNTNWCAGEGAGSCGDSGGLVCGGILLOGIYSWMD  
 VPDDNTKKPGVYTKVCHYLEMIRETKMRN"  
 <42595..>47769  
 /gene="KLK3"  
 /note="serine protease"  
 /codon\_start=1  
 /product="prostate specific antigen"  
 join(<42595..42640,43880..44039,45669..45955,46099..46235,  
 47614..47769)  
 /gene="KLK3"  
 /product="prostate specific antigen"  
 join(<42595..42640,43880..44039,45669..45955,46099..46235,  
 47614..47769)  
 /gene="KLK3"  
 /note="serine protease"  
 /codon\_start=1  
 /product="prostate specific antigen"  
 /protein\_id="AAG33355.1"  
 /db\_xref="GI:11244760"  
 /translation="MWVYVFLVLTWTIGAAPLILSRIVGWECEKHSOPWQVLVAVS  
 RGRVACGGVAVHPQWLTAAHCIRKNSVILGRSLFHPEDTGVQFVSHSPHPLVD  
 MSLIKNRFLEPPDSDSHDMLRLSEPAELTDAYKVDLPLOEPALGTTCYASGWSGI  
 EPEEFLTPKKLCVDLHVISNDVCAQVHPQKVKPKMDCAGRMGKSTCGSDSGPLV  
 CNGVLOGITWSGSPCALPERSPLTYVVKVKKIKDTIVANP"  
 <61139..>66229  
 /gene="KLK2"  
 /note="serine protease"  
 /codon\_start=1  
 /product="glandular kallikrein"  
 join(<61139..61184,62391..62550,64142..64428,64542..64678,  
 66074..66229)  
 /gene="KLK2"  
 /product="glandular kallikrein"  
 join(<61139..61184,62391..62550,64142..64428,64542..64678,  
 66074..66229)  
 /gene="KLK2"  
 /note="serine protease"  
 /codon\_start=1  
 /product="glandular kallikrein"  
 /protein\_id="AAG33356.1"  
 /db\_xref="GI:11244761"  
 /translation="MMDLVLSIALSVGTGAVPLIOSRIVGWECEKHSOPWQVLVAVS  
 HGNVACGGVAVHPQWLTAAHCIRKNSVILGRSLFHPEDTGVQFVSHSPHPLVD  
 MSLIKNRFLEPPDSDSHDMLRLSEPAELTDAYKVDLPLOEPALGTTCYASGWSGI  
 EPEEFLTPKKLCVDLHVISNDVCAQVHPQKVKPKMDCAGRMGKSTCGSDSGPLV  
 CNGVLOGITWSGSPCALPERSPLTYVVKVKKIKDTIVANP"  
 <61139..>66229  
 /gene="KLK2"  
 /note="serine protease"  
 /codon\_start=1  
 /product="glandular kallikrein"  
 /protein\_id="AAG33357.1"  
 /db\_xref="GI:11244762"  
 /translation="MATAGNMGWFLGTLIGVAVSVSGSQIINGECPSPHSPQ  
 OALVYENELFCGSLVYHPQWVLSAAHCIRKNSVILGRSLFHPEDTGVQFVSHSPHPLVD  
 VRIPEKRLPLANDMLIKLDESIVSESDTIRTSIAQCPTAGNSCLVSGMGLANR  
 MPTVLOQVSVSVSEVCSKLYDPLVHPNMFCAAGGDDOCSGDSGGLVCGGILLOGIYSWMD  
 GLVSGFAPCGGVPGVYTNLCFTFEMIEKTVQAS"  
 complement(<131301..>140303)  
 /gene="KLK5"  
 /note="synonym: SCPE"  
 complement(join(<131301..131456,136110..136443,  
 136529..136785,137525..137690,140255..>140303))  
 /gene="KLK5"  
 /product="stratum corneum trypsin-like serine protease"  
 complement(join(131301..131456,136310..136443,  
 136529..136785,137525..137690,140255..140303))  
 /gene="KLK5"  
 /codon\_start=1  
 /product="stratum corneum trypsin-like serine protease"  
 /protein\_id="AAG33358.1"  
 /db\_xref="GI:11244763"

/translation="MWVLCALITALLIGVADARSDSSRTINGSDCMHTPQMOAL  
 LTRNPQVCGAVLVHPQWLTAAHCIRKNSVILGRSLFHPEDTGVQFVSHSPHPLVD  
 PGVSHPGNSNDMLIKLNRIRPTKDRPPLNVSHSCASACTGLVSGWGTGSPQVHE  
 PKYLOCLNIVLSQKRCEDAVPQIDTMCAGKARDCSDGSGPVVVCNSLQGL  
 VSMEDYCARPNRGVYTNLCFTFEMIEKTVQAS"  
 complement(join(<146834..146986,149628..149764,  
 151186..151433,155052..155208,155948..>155987))  
 /product="protease M"  
 complement(join(146834..146986,149628..149764,  
 151186..151433,155052..155208,155948..155987))  
 /note="serine protease; also called neurosin or zyme"  
 /codon\_start=1  
 /product="protease M"  
 /protein\_id="AAG33359.1"  
 /db\_xref="GI:11244764"  
 /translation="MKKLVAVLSIAAAMAEQNKLVHGSPCDKTSHPYQALYTSCH  
 LDCGVLIHPLVLTAAHCIRKNSVILGRSLFHPEDTGVQFVSHSPHPLVD  
 HDQIMLRAPAKLSELIOPLREDCSANTTSCHILGKGTADGDPDTIOCAVI  
 HLYSRECEHAYRGQITQNMLCAGDEKYGKSDSGGLVCGDHLRLGLVSGNTPC  
 GSKERPGVYTNVCTNWTMTQITQAK  
 complement(<165420..>170283)  
 /gene="KLK7"  
 /note="synonym: SCPE"  
 complement(join(<165420..165575,167672..167808,  
 168124..168371,169651..169798,170211..>170283))  
 /gene="KLK7"  
 /product="stratum corneum chymotryptic enzyme"  
 complement(join(165420..165575,167672..167808,  
 168124..168371,169651..169798,170211..170283))  
 /gene="KLK7"  
 /note="serine protease"  
 /codon\_start=1  
 /product="stratum corneum chymotryptic enzyme"  
 /protein\_id="AAG33360.1"  
 /db\_xref="GI:11244765"  
 /translation="MARSLLPQIILLSALFAGEAOGDKTIQDAPCARSHPMQ  
 VALISGNOLHCGVLYNERVLTAAHCIRKNSVILGRSLFHPEDTGVQFVSHSPHPLVD  
 GYSTQHVNDMLVRLKNSQRLSSVAKVPLPSCPEPGTTCYSGGTTSPDVTPE  
 VSMGTEPCGPNPGVYTNLCFTFEMIEKTVQAS"  
 complement(join(<183943..184098,185635..185768,  
 187865..188127,188293..188452,188967..>189036))  
 /product="neurosin"  
 complement(join(183943..184098,185635..185768,  
 187865..188127,188293..188452,188967..189036))  
 /note="serine protease; also called ovasin"  
 /codon\_start=1  
 /product="neurosin"  
 /protein\_id="AAG33361.1"  
 /db\_xref="GI:11244766"  
 /translation="MGRPRRAKTYMFLLLLGAMAGHRAQEDKVLGGHECOPHSQ  
 PMQALFQGOGLCGVLYGCGNWLTPAAHCIRKNSVILGRSLFHPEDTGVQFVSHSPHPLVD  
 SIPHCYNSDVEDHNDMLLORDASGSKRPISLDHCTOPKCTVSGMGTV  
 TSPRENPDITNCAEVKIFPQKCEADYQSLQIDGMVCAAGDAPCQDSDSGPLVC  
 DGLQGTSGSPGCSKDPGYVNTICRILDMTKIKSGK"  
 complement(join(<190980..191129,191573..191709,  
 194324..194589,197048..197204,197370..>197412))  
 /product="kallikrein-like 3"  
 complement(join(190980..191129,191573..191709,  
 194324..194589,197048..197204,197370..197412))  
 /note="serine protease"  
 /codon\_start=1  
 /product="kallikrein-like 3"  
 /protein\_id="AAG33362.1"  
 /db\_xref="GI:11244767"  
 /translation="MKLGLCALLSLAGHMDTRTGAEGCRPNQSPQAGLPHLT  
 RLFCGATLISDRMLTAAHCIRKNSVILGRSLFHPEDTGVQFVSHSPHPLVD  
 LLSANDHNDIMLRPLRQARLSAVQYLNLSQTCVSGFGMGLCLSGKAVSSPALPVP  
 TLQCCASILENKLCHWALPGHISDMLCAGLWGGGSCQSDSGGLVCGDHLRLGLVSGNTPC  
 SGGAEPCSRPRRAVYTVSYCHYLDWIDQIEIMEN"

Query Match 100.0%; Score 2071; DB 9; Length 230000;  
 Best Local Similarity 100.0%; Pred. No. 0;



```

DEFINITION Homo sapiens chromosome 19 clone.RP11-795B6 map 19, WORKING DRAFT
SEQUENCE 12 unordered pieces.
ACCESSION AC027602
VERSION AC027602.4 GI:11178143
KEYWORDS HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
TITLE Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
JOURNAL Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G.,
REFERENCE Campoliano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
AUTHORS Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marguis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
McDermid,J., Menues,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., O'Neil,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Triggilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
JOURNAL Direct Submission
TITLE Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 217346)

Direct Submission
TITLE Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS Smit, A.F.A. & Green, P. (1996-1997)
COMMENT All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9166

```

```

----- Summary Statistics
Center clone name: 795.B.6
Sequencing vector: M13, M7815; 31% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 210748 bases at least Q40
Consensus quality: 215058 bases at least Q30
Insert size: 194000; agarose-fp
Quality coverage: 11.9 in Q20
Quality coverage: 10.7 in Q20.
NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
8149: contig of 8149 bp in length
8150: gap of 100 bp
8250: contig of 1343 bp in length
9592: gap of 100 bp
9593: gap of 100 bp
9593: contig of 1041 bp in length
10734: gap of 100 bp
10833: gap of 100 bp
10834: contig of 2686 bp in length
13520: gap of 100 bp
13619: gap of 100 bp
13620: contig of 3891 bp in length
17510: gap of 100 bp
17511: gap of 100 bp
17610: contig of 6992 bp in length
17611: gap of 100 bp
24602: contig of 10732 bp in length
24702: gap of 100 bp
24703: gap of 100 bp
35434: contig of 10732 bp in length
35435: gap of 100 bp
35435: gap of 100 bp
124474: contig of 88940 bp in length
124475: gap of 100 bp
124575: gap of 100 bp
134664: contig of 10090 bp in length
134665: gap of 100 bp
134764: gap of 100 bp
162343: contig of 27579 bp in length
162344: gap of 100 bp
162444: gap of 100 bp
208917: contig of 46474 bp in length
208918: gap of 100 bp
209017: gap of 100 bp
217346: contig of 8329 bp in length.
Location/Qualifiers
1. 217346
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19"
/clone="RP11-795B6"
/clone_1lb="RP11-11 Human Male BAC"
1. 8149
/clone="assembly_fragment"
vector_end:Sp6
vector_side:left"
8250..9592
/clone="assembly_fragment"
9693..10733
/clone="assembly_fragment"
10834..13519
/clone="assembly_fragment"
13620..17510
/clone="assembly_fragment"
17611..24602
/clone="assembly_fragment"
24703..35434
/clone="assembly_fragment"
35535..124474
/clone="assembly_fragment"
124575..134664
/clone="assembly_fragment"

```

```
misc_feature 134765..162343
                /note="assembly-fragment"
misc_feature 162444..208917
                /note="assembly-fragment"
misc_feature 209018..217346
                /note="assembly-fragment"
                clone_end:77
                vector_side:right"
BASE COUNT 55109 a 54154 c 53923 g 53053 t 1107 others
ORIGIN
Query Match 96.1%; Score 1990; DB 2; Length 217346;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 2045; Conservative 0; Mismatches 5; Indels 22; Gaps 3;
QY 1 AAACAGACCAAAACCTCCGCGCGGTGAGACCTCATGTTCCCAAGTGGAAAGCAGAGCA 60
Db 194902 AAACAGACCAAAACCTCCGCGCGGTGAGACCTCATGTTCCCAAGTGGAAAGCAGAGCA 194843
QY 61 TAAAGATAGATTAATATGTAGTAATTAATAAAAAAAAAAAAAAAAAATAGCCGGGTGTGTG 120
Db 194842 TAAAGATAGATTAATATGTAGTAATTAATAAAAAAAAAAAAAAAAAATAGCCGGGTGTGTG 194785
QY 121 CTTGACCTGTAGTTCAGCTACTTGGAGGCTGAGAGTGGGAGAAATTGCTTGAGCCCAA 180
Db 194784 CTTGACCTGTAGTTCAGCTACTTGGAGGCTGAGAGTGGGAGAAATTGCTTGAGCCCAA 194725
QY 181 CGTTTGAAGCTCGGTGAGACCATGACCTGCTGCTCCAGACAGCAGCTGGGTGA 240
Db 194724 CGTTTGAAGCTCGGTGAGACCATGACCTGCTGCTCCAGACAGCAGCTGGGTGA 194665
QY 241 CAAGCAAGACGTTTTTGTGAGAAAGAAAAAGAGAGAGAGAGAGAGAGAGAG 300
Db 194664 CAAGCAAGACGTTTTTGTGAGAAAGAAAAAGAGAGAGAGAGAGAGAGAGAG 194605
QY 301 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 194604 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 194564
QY 361 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 194563 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 194504
QY 421 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 194503 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 194444
QY 481 TATTATAGAGAGGTATTTATAGGAGGTATGGGATTTGAAGCAGAGAAACAAATTAG 540
Db 194443 TATTATAGAGAGGTATTTATAGGAGGTATGGGATTTGAAGCAGAGAAACAAATTAG 194384
QY 541 TCCAGCCGATGATTTCTATTTGGAGTATTTGCCCCAGAGAGAGAGAGAGAG 600
Db 194383 TCCAGCCGATGATTTCTATTTGGAGTATTTGCCCCAGAGAGAGAGAGAGAG 194324
QY 601 GAGAGATTTTGTGTGTGACAACTATATGAGAGGAGCACTTATGCACTATGATAG 660
Db 194323 GAGAGATTTTGTGTGTGACAACTATATGAGAGGAGCACTTATGCACTATGATAG 194264
QY 661 TGCCAGTGTGCTTCAACATGCTATGATGACACAGGAGGCTCCACAACCAACATT 720
Db 194263 TGCCAGTGTGCTTCAACATGCTATGATGACACAGGAGGCTCCACAACCAACATT 194204
QY 721 ATCCAGCTTCAGATGCCACAGTGGCCAGATCGAGAGAACCTTCATCCAGGGCTGAGAAC 780
Db 194203 ATCCAGCTTCAGATGCCACAGTGGCCAGATCGAGAGAACCTTCATCCAGGGCTGAGAAC 194144
QY 781 CGTATTTTTCAGAGAGAGATTAAGAGTGGGTGTGTGAGAGATGGGAGAGAGAGGTGT 840
Db 194143 CGTATTTTTCAGAGAGAGATTAAGAGTGGGTGTGTGAGAGATGGGAGAGAGAGGTGT 194084
QY 841 GTGTCCAGTAAAGAAATTAAGGCTTCACAGCTGAGAGGAGAGAGAGAGAGAGAG 900
|||||
```

```
Db 194083 GTGTCCAGTAAAGAAATTAAGGCTTCACAGCTGAGAGGAGAGAGAGAGAGAG 194024
QY 901 GCGGAGAGATACACGATGAGGAGAGAGAGCTGGAACAGAAAGTAGAGAGAGAGAGAG 960
Db 194023 GCGGAGAGATACACGATGAGGAGAGAGAGAGCTGGAACAGAAAGTAGAGAGAGAGAG 193964
QY 961 GATGTGAGAGAGAGAGGTCACAGAGACCCCGCAATATGATGTGTGACAAACAGAAATCTTG 1020
Db 193963 GATGTGAGAGAGAGAGGTCACAGAGACCCCGCAATATGATGTGTGACAAACAGAAATCTTG 193904
QY 1021 AAGAGAAATATGAGTGGAGAGATGACAAATGGGGTCTTAAGTGTAACTTGAAGCCAGAG 1080
Db 193903 AAGAGAAATATGAGTGGAGAGATGACAAATGGGGTCTTAAGTGTAACTTGAAGCCAGAG 193844
QY 1081 CATGGTGTCAAGCCCTGTAAATCCCAACCTTTGGAGGTGAGTGGGAGCAATCTTGA 1140
Db 193843 CATGGTGTCAAGCCCTGTAAATCCCAACCTTTGGAGGTGAGTGGGAGCAATCTTGA 193784
QY 1141 GGCACAGAGTTGAGACACAGCCCTGGCCAAATGTTGAAACCCGTCCTTACAAAAAAT 1200
Db 193783 GGCACAGAGTTGAGACACAGCCCTGGCCAAATGTTGAAACCCGTCCTTACAAAAAAT 193724
QY 1201 ACAAAAATTTAGCCGGGTGTGGTGAATGAGACCTGTAGTACAGCTACTTGGAGGCTGA 1260
Db 193723 ACAAAAATTTAGCCGGGTGTGGTGAATGAGACCTGTAGTACAGCTACTTGGAGGCTGA 193664
QY 1261 GGCAGAGAGATTTGTTGAACCCCGGAGATGAGAGGCTGACAGTGAAGTCAAGGCTCACT 1320
Db 193663 GGCAGAGAGATTTGTTGAACCCCGGAGATGAGAGGCTGACAGTGAAGTCAAGGCTCACT 193604
QY 1321 GCGCTCAACCTGGGACACAGAGTAAAGTCACTCTC -AAAAAAAAAAGCTGGATTGG 1379
Db 193603 GCGCTCAACCTGGGACACAGAGTAAAGTCACTCTC -AAAAAAAAAAGCTGGATTGG 193544
QY 1380 GAGTGAATATTAATTAATTAATTTCCCTCTCTCCCTTTGGCTGGTGTCTCCTCTGTC 1439
Db 193543 GAGTGAATATTAATTAATTAATTTCCCTCTCTCTCTCTCTCTTTGCTGTCTCTCTCTGTC 193484
QY 1440 TTTTCTGATTTCTTCATCTCTGTAATCTTTCATCTCTGTGTCTGTCTCCATCTGCTT 1499
Db 193483 TTTTCTGATTTCTTCATCTCTGTAATCTTTCATCTCTGTGTCTGTCTCCATCTGCTT 193424
QY 1500 CTCATCTATGGGATCTCTGGGTCTCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1559
Db 193423 CTCATCTATGGGATCTCTGGGTCTCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 193364
QY 1560 TGCCTCTCATAGCCCCCTTCTCTCTCTGAGAGGAGATTCGGGGGCTCTGGTCTGTC 1619
Db 193363 TGCCTCTCTCTGCCCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGTC 193304
QY 1620 AATGGCTCCCTCAGAGGACTGTGTCTCTGAGAGATTAATCTTGTGCCCCGCCCAACAGA 1679
Db 193303 AATGGCTCCCTCAGAGGACTGTGTCTCTGAGAGATTAATCTTGTGCCCCGCCCAACAGA 193244
QY 1680 CCGGGTGTCTACACGAACTCTGCAAGTTCCACCAAGTGGATCCAGAAACATCCAGGCTC 1739
Db 193243 CCGGGTGTCTACACGAACTCTGCAAGTTCCACCAAGTGGATCCAGAAACATCCAGGCTC 193184
QY 1740 AACTCTGATCATCCAGAGACTCAGACACAGCGGATCCCACTCTGCTCAGAGAGAGGCT 1799
Db 193183 AACTCTGATCATCCAGAGACTCAGACACAGCGGATCCCACTCTGCTCAGAGAGAGGCT 193124
QY 1800 CTGACATCTCTTTCAGACCCCTCATCTCTTCCAGAGATTTGAGATGTTCTCTCTCA 1859
Db 193123 CTGACATCTCTTTCAGACCCCTCATCTCTTCCAGAGATTTGAGATGTTCTCTCTCA 193064
QY 1860 GCGCCTGACCCATGCTCTCTGAGACTCAGGGTCTGCTTCCCACTTGGGCTGACCGTG 1919
Db 193063 GCGCCTGACCCATGCTCTCTGAGACTCAGGGTCTGCTTCCCACTTGGGCTGACCGTG 193004
QY 1920 TCTCTCTAGTTGAACCTGGGAAATTTCCAAACTGTCAGAGGCGGGGCTTGGGCTCTC 1979
Db 193003 TCTCTCTAGTTGAACCTGGGAAATTTCCAAACTGTCAGAGGCGGGGCTTGGGCTCTC 192944
```

Db	Accession	Definition	Accession	Definition
Y	1980		AC130782	200792 bp DNA 11near HTG 22-NOV-2002
Y	1980		AC130782	ordered pieces.
Y	192943		AC130782	AC130782.2 GI:25167101
Y	192943		HTG: HTGS_PHASE2: HTGS_DRAFT.	
Y	192943		Pan troglodytes (chimpanzee)	
Y	192943		Pan troglodytes	
Y	192943		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Y	192943		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.	
Y	192943		1 (bases 1 to 200792)	
Y	192943		Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,	
Y	192943		Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,	
Y	192943		Carlae,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,	
Y	192943		Haghighi,P., Han,J., Hansen,N., Ho,S.L., Idol,J.R., Karlins,E.,	
Y	192943		Latic,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,O.L., Maduro,V.B.,	
Y	192943		Margulies,E.H., Masiello,C., Maskeri,B., McDowell,D.,	
Y	192943		Paquirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,	
Y	192943		Redix-Duque,N., Schandler,K., Schueler,M.G., Sison,C.,	
Y	192943		Stentipop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,	
Y	192943		Weherby,K.D., Wiggins,L., Young,A. and Green,E.D.	
Y	192943		NISC Comparative Sequencing Initiative	
Y	192943		Unpublished	
Y	192943		2 (bases 1 to 200792)	
Y	192943		Green,E.D.	
Y	192943		Direct Submission	
Y	192943		Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717	
Y	192943		Groventom Circle, Gaithersburg, MD 20877, USA	
Y	192943		3 (bases 1 to 200792)	
Y	192943		Green,E.D.	
Y	192943		Direct Submission	
Y	192943		Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717	
Y	192943		Groventom Circle, Gaithersburg, MD 20877, USA	
Y	192943		On Nov 22, 2002 this sequence version replaced gi:22218452.	
Y	192943		Genome Center	
Y	192943		Center: NIH Intramural Sequencing Center	
Y	192943		Center code: NISC	
Y	192943		Web site: http://www.nisc.nih.gov	
Y	192943		Contact: nisc.zoo@nih.gov	
Y	192943		Project Information	
Y	192943		Center project name: dHz	
Y	192943		Center clone name: 355A20	
Y	192943		Summary Statistics	
Y	192943		Sequencing vector: plasmid; n/a; 100% of reads	
Y	192943		Chemistry: Dye-terminator Big Dye; 100% of reads	
Y	192943		Assembly program: Phrap; version 0.990319	
Y	192943		Consensus quality: 19656 bases at least Q40	
Y	192943		Consensus quality: 197883 bases at least Q50	

```

Consensus quality: 198879 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 199692; sum-of-contigs
Quality coverage: 9.03x in Q20 bases; agarose-fp
Quality coverage: 9.73x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 28306: contig of 28306 bp in length
* 28406: gap of unknown length
* 32807 37856: contig of 9450 bp in length
* 37857 37956: gap of unknown length
* 37957 73522: contig of 35566 bp in length
* 73523 73623: gap of unknown length
* 73623 83567: contig of 9945 bp in length
* 83568 83667: gap of unknown length
* 83668 88817: contig of 5150 bp in length
* 88818 88917: gap of unknown length
* 88918 125611: contig of 36594 bp in length
* 125612 125711: gap of unknown length
* 125712 159879: contig of 34168 bp in length
* 159880 159979: gap of unknown length
* 159980 174698: contig of 14719 bp in length
* 174699 174798: gap of unknown length
* 174799 186382: contig of 11584 bp in length
* 186383 186482: gap of unknown length
* 186483 193344: contig of 6862 bp in length
* 193345 193444: gap of unknown length
* 193445 199463: contig of 5919 bp in length
* 199464 199463: gap of unknown length
* 199464 200792: contig of 1329 bp in length.
Location/Qualifiers
FEATURES
source
1. 200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-355A20"
/clone_1lb="CH251"
1. 28306
/note="assembly_fragment"
vector_end:Sp6
vector_side:left"
misc_feature
28407..37856
/note="assembly_fragment"
misc_feature
37957..73522
/note="assembly_fragment"
misc_feature
73623..83567
/note="assembly_fragment"
misc_feature
83668..88817
/note="assembly_fragment"
misc_feature
88918..125611
/note="assembly_fragment"
misc_feature
125712..159879
/note="assembly_fragment"
misc_feature
159980..174698
/note="assembly_fragment"
misc_feature
174799..186382
/note="assembly_fragment"
misc_feature
186483..193344
/note="assembly_fragment"
misc_feature
193445..199363
/note="assembly_fragment"
misc_feature
199464..200792
/note="assembly_fragment"
vector_end:T7
vector_side:right"

```



```

RESULT 6
AX331407/c
LOCUS AX331407 586 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 1916 from Patent WO0194629.
ACCESSION AX331407
VERSION AX331407.1 GI:18122041
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
          Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
          gene sets
JOURNAL Patent: WO 0194629-A 1916 13-DEC-2001;
          Avalon Pharmaceuticals (US)
FEATURES
Source location/Qualifiers
BASE COUNT 130 a 139 c 179 g 136 t 2 others
ORIGIN
Query Match 23.3%; Score 482; DB 6; Length 586;
Best Local Similarity 100.0%; Pred. No. 8.1e-111;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGATGATTCCTGGGGGCGCTGTGTCGATGCTCCCTCGCAGGAGCTGTCTCTGG 1649
    |||||||
DB 501 CAGGGATGATTCCTGGGGGCGCTGTGTCGATGCTCCCTCGCAGGAGCTGTCTCTGG 442
    |||||||

QY 1650 GGAGATTACCTTGTGGCGCGCCCAACAGACGGGGGTGTACAGAACTCTGCAAGTTTC 1709
    |||||||
DB 441 GGAGATTACCTTGTGGCGCGCCCAACAGACGGGGGTGTACAGAACTCTGCAAGTTTC 382
    |||||||

QY 1710 ACCAAGTGATTCAGAGAAACCATCCAGGCACTCTGAGTCAATCCAGAGCTCAGACACA 1769
    |||||||
DB 381 ACCAAGTGATTCAGAGAAACCATCCAGGCACTCTGAGTCAATCCAGAGCTCAGACACA 322
    |||||||

QY 1770 CCGGCATCCCGACCTGTGTGAGGAGACAGCCCTGACACTCTTTCAAGACCTCATTCCTTC 1829
    |||||||
DB 321 CCGGCATCCCGACCTGTGTGAGGAGACAGCCCTGACACTCTTTCAAGACCTCATTCCTTC 262
    |||||||

QY 1830 CCAGAGATGTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCCTGAGACTCAGG 1889
    |||||||
DB 261 CCAGAGATGTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCCTGAGACTCAGG 202
    |||||||

QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGGAACAATTTTC 1949
    |||||||
DB 201 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGGAACAATTTTC 142
    |||||||

QY 1950 CAAAGTGTCCAGAGGGGGGGGTGCGTCTCAATCTCCCTGGGACATTTATCTCAAGC 2009
    |||||||
DB 141 CAAAGTGTCCAGAGGGGGGGGTGCGTCTCAATCTCCCTGGGACATTTATCTCAAGC 82
    |||||||

QY 2010 TCAGGGCCCATCCCTTCTCTGACAGCTGTGACCCAAATTTAGTCCAGAAATAAAGTGA 2069
    |||||||
DB 81 TCAGGGCCCATCCCTTCTCTGACAGCTGTGACCCAAATTTAGTCCAGAAATAAAGTGA 22
    |||||||

QY 2070 AG 2071
    ||
DB 21 AG 20

```

```

RESULT 7
AX331824/c
LOCUS AX331824 586 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 2333 from Patent WO0194629.

```

```

ACCESSION AX331824
VERSION AX331824.1 GI:18122458
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
          Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
          gene sets
JOURNAL Patent: WO 0194629-A 2333 13-DEC-2001;
          Avalon Pharmaceuticals (US)
FEATURES
Source location/Qualifiers
BASE COUNT 130 a 139 c 179 g 136 t 2 others
ORIGIN
Query Match 23.3%; Score 482; DB 6; Length 586;
Best Local Similarity 100.0%; Pred. No. 8.1e-111;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGATGATTCCTGGGGGCGCTGTGTCGATGCTCCCTCGCAGGAGCTGTCTCTGG 1649
    |||||||
DB 501 CAGGGATGATTCCTGGGGGCGCTGTGTCGATGCTCCCTCGCAGGAGCTGTCTCTGG 442
    |||||||

QY 1650 GGAGATTACCTTGTGGCGCGCCCAACAGACGGGGGTGTACAGAACTCTGCAAGTTTC 1709
    |||||||
DB 441 GGAGATTACCTTGTGGCGCGCCCAACAGACGGGGGTGTACAGAACTCTGCAAGTTTC 382
    |||||||

QY 1710 ACCAAGTGATTCAGAGAAACCATCCAGGCACTCTGAGTCAATCCAGAGCTCAGACACA 1769
    |||||||
DB 381 ACCAAGTGATTCAGAGAAACCATCCAGGCACTCTGAGTCAATCCAGAGCTCAGACACA 322
    |||||||

QY 1770 CCGGCATCCCGACCTGTGTGAGGAGACAGCCCTGACACTCTTTCAAGACCTCATTCCTTC 1829
    |||||||
DB 321 CCGGCATCCCGACCTGTGTGAGGAGACAGCCCTGACACTCTTTCAAGACCTCATTCCTTC 262
    |||||||

QY 1830 CCAGAGATGTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCCTGAGACTCAGG 1889
    |||||||
DB 261 CCAGAGATGTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCCTGAGACTCAGG 202
    |||||||

QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGGAACAATTTTC 1949
    |||||||
DB 201 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGGAACAATTTTC 142
    |||||||

QY 1950 CAAAGTGTCCAGAGGGGGGGGTGCGTCTCAATCTCCCTGGGACATTTATCTCAAGC 2009
    |||||||
DB 141 CAAAGTGTCCAGAGGGGGGGGTGCGTCTCAATCTCCCTGGGACATTTATCTCAAGC 82
    |||||||

QY 2010 TCAGGGCCCATCCCTTCTCTGACAGCTGTGACCCAAATTTAGTCCAGAAATAAAGTGA 2069
    |||||||
DB 81 TCAGGGCCCATCCCTTCTCTGACAGCTGTGACCCAAATTTAGTCCAGAAATAAAGTGA 22
    |||||||

QY 2070 AG 2071
    ||
DB 21 AG 20

```

```

RESULT 8
G41906/c
LOCUS G41906 586 bp DNA linear STS 30-SEP-1998
DEFINITION SHGC-56840 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION G41906
VERSION G41906.1 GI:3668239
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



REFERENCE 1 (bases 1 to 586)  
 AUTHORS Myers R.M.  
 TITLE Human STS (1998)  
 JOURNAL Unpublished (1998)  
 COMMENT

Contact: Richard M. Myers  
 Stanford Human Genome Center (SHGC)  
 Stanford University School of Medicine  
 Department of Genetics, M-344, Stanford, CA 94305, USA  
 Tel: 4157259687  
 Fax: 4157259689  
 Email: myers@shgc.stanford.edu  
 Primer A: AGACACAGGTCAGCCCAAT  
 Primer B: GCCAATCTGAGTCATCCC  
 STS size: 188  
 PCR profile:

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 30 seconds

Annealing: 60 degrees C for 30 seconds

Polymerization: 72 degrees C for 23 seconds

PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9700

Protocol:

Template: 25 ng

Primer: each 1 uM

dNTPs: each 200 uM

Amplifrag Gold Polymerase: 0.07 units/ul

Total Vol: 5 ul

Buffer:

MgCl2: 2.5 mM

KCl: 50 mM

Tris-HCl: 10 mM

pH: 8.3

Prepared with primer pairs derived from W73140 -- Unigene.

Location/Qualifiers

1..586

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/map="19"

/clone\_id="Human"

STS

primer\_bind

primer\_bind

BASE COUNT

130 a 179 g 136 t 2 others

Query Match

Best Local Similarity 100.0%; Pred. No. 8.1e-11; Length 586;

Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 261 CCAGAGATGTTGAGATGTTGATCTCTCCAGCCCCCTGACCCCATGCTCTGAGCTGAG 202

Qy 1890 GTCTGCTTCCCGCCACATTTGGGCTGACGCTGCTCTCTGATGTAACCCCGGAGCAATTTG 1949

Db 201 GTCTGCTTCCCGCCACATTTGGGCTGACGCTGCTCTCTGATGTAACCCCGGAGCAATTTG 142

Qy 1950 CAAACTGTCCAGAGGCGGGGCTTGGCTCAATCTCCCTGGGGCACTTTCATCTCAAGC 2009

Db 141 CAAACTGTCCAGAGGCGGGGCTTGGCTCAATCTCCCTGGGGCACTTTCATCTCAAGC 82

Qy 2010 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTATCCCAAGAAATTAAGTGA 2069

Db 81 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTATCCCAAGAAATTAAGTGA 22

Qy 2070 AG 2071

Db 21 AG 20

RESULT 9

AX429955/c 735 bp DNA Linear PAT 21-JUN-2002

LOCUS AX429955

DEFINITION Sequence 94 from Patent WO0198339.

ACCESSION AX429955

VERSION AX429955.1 GI:21541119

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.

Compositions and methods for the therapy and diagnosis of breast

cancer

Patent: WO 0198339-A 94 27-DEC-2001;

CORTIXA CORPORATION (US)

FEATURES

Location/Qualifiers

1..735

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT

161 a 176 c 226 g 172 t

Query Match

Best Local Similarity 100.0%; Pred. No. 8.3e-11; Length 735;

Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1590 CAGGGTGAATTCGGGGGCGCTGTGTGTCGCAATGGCTCCCTGACAGGACTGTCCTGG 1649

Db 500 CAGGGTGAATTCGGGGGCGCTGTGTGTCGCAATGGCTCCCTGACAGGACTGTCCTGG 441

Qy 1650 GGAGATTACCTTGTGCCCCGCCCAACAGACCGGGTGTCTACAGAACCTTGCAGATTG 1709

Db 440 GGAGATTACCTTGTGCCCCGCCCAACAGACCGGGTGTCTACAGAACCTTGCAGATTG 381

Qy 1710 ACCAAGTGAATTCAGGAACCAATCCAGGCCAATCTAGTCATGCCAGACTGAGCACA 1769

Db 380 ACCAAGTGAATTCAGGAACCAATCCAGGCCAATCTAGTCATGCCAGACTGAGCACA 321

Qy 1770 CGGGATCCCACTGCTGTCAGAGGACAGCCCTGACATCTTTCAGAGCCCTCATCTTC 1829

Db 320 CGGGATCCCACTGCTGTCAGAGGACAGCCCTGACATCTTTCAGAGCCCTCATCTTC 261

Qy 1830 CCAGAGATGTTGAGATGTTGATCTCTCCAGCCCCCTGACCCCATGCTCTGAGCTGAG 1889

Db 260 CCAGAGATGTTGAGATGTTGATCTCTCCAGCCCCCTGACCCCATGCTCTGAGCTGAG 201

Qy 1890 GTCTGCTTCCCGCCACATTTGGGCTGACGCTGCTCTCTGATGTAACCCCGGAGCAATTTG 1949

Db 200 GTCTGCTTCCCGCCACATTTGGGCTGACGCTGCTCTCTGATGTAACCCCGGAGCAATTTG 141

Qy 1950 CAAACTGTCCAGAGGCGGGGCTTGGCTCAATCTCCCTGGGGCACTTTCATCTCAAGC 2009



|||||  
Db 140 CAAAGTGTCCAGGGGGGGGGTGGCTCAATCTCCCTGGGGGACATTTATCTCAAGC 81  
QY 2010 TCAGGGCCCATCCCTTCTCTGACGCTGTGACCCAAATTTAGTCCAGAAATAACTGAGA 2069  
Db 80 TCAGGGCCCATCCCTTCTCTGACGCTGTGACCCAAATTTAGTCCAGAAATAACTGAGA 21  
QY 2070 AG 2071  
||  
Db 20 AG 19  
RESULT 10  
BD139877/c 738 bp DNA linear PAT 18-SEP-2002  
LOCUS Compounds for immunotherapy and diagnosis of breast cancer and  
DEFINITION methods for their use.  
ACCESSION BD139877  
VERSION BD139877.1 GI:23234822  
KEYWORDS JP 2002507387-A/94.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 738)  
AUTHORS Reed,S.G. and Xu,J.  
TITLE Compounds for immunotherapy and diagnosis of breast cancer and  
METHODS for their use  
PATENT: JP 2002507387-A 94 12-MAR-2002;  
JOURNAL CORIXA CORP  
OS Homo sapiens (human)  
PN JP 2002507387-A/94  
PD 12-MAR-2002  
PR 22-DEC-1998 JP 2000526543  
PR 24-DEC-1997 US 08/998253,24-DEC-1997 US 08/998255 PR  
17-JUL-1998 US 09/118627,17-JUL-1998 US 09/118554 PI STEVEN  
G REED,JIANGCHUN XU  
PC C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,  
PC C07K14/47,  
PC C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/PC  
53,  
PC G01N33/577,C12N15/00,A61K37/02,C12N5/00  
CC Compounds for immunotherapy and diagnosis  
of breast cancer and  
CC their use methods for  
CC key Location/Qualifiers  
FH source 1..738  
FT Location/Qualifiers  
1..738  
source /organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 161 a 179 c 226 g 172 t  
ORIGIN  
Query Match 23.3%; Score 482; DB 6; Length 738;  
Best Local Similarity 100.0%; Pred. No. 8.3e-111;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1590 CAGGGATTCCTGGGGGGGGCTGTGCTGATGAGTCCCTGCAGGAGCTGCTGCTGG 1649  
Db 503 CAGGGATTCCTGGGGGGGGCTGTGCTGATGAGTCCCTGCAGGAGCTGCTGCTGG 444  
QY 1650 GAGATTACCTGTGCTGCCCGCCCAACAGACCGGCTGTACAGAGACCTCTGCAAGTTC 1709  
Db 443 GAGATTACCTGTGCTGCCCGCCCAACAGACCGGCTGTACAGAGACCTCTGCAAGTTC 384  
QY 1710 ACCAAGTGATCCAGAAACCATTCAGGCCCACTCTGAGTCAATCCAGAGCTCAGACACA 1769  
Db 383 ACCAAGTGATCCAGAAACCATTCAGGCCCACTCTGAGTCAATCCAGAGCTCAGACACA 324

QY 1770 CCGGCATCCCGACCTGTGACGGAGACGCCCTGACACTCTCTTCAGACCTCATTCCTTC 1829  
Db 323 CCGGCATCCCGACCTGTGACGGAGACGCCCTGACACTCTCTTCAGACCTCATTCCTTC 264  
QY 1830 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATCTCTCTGAGCTCAG 1889  
Db 263 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATCTCTCTGAGCTCAG 204  
QY 1890 GTCTGCTCCCGACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGGGAAATTTTC 1949  
Db 203 GTCTGCTCCCGACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGGGAAATTTTC 144  
QY 1950 CAAAGTGTCCAGGGGGGGGGTGGCTCAATCTCCCTGGGGGACATTTATCTCAAGC 2009  
Db 143 CAAAGTGTCCAGGGGGGGGGTGGCTCAATCTCCCTGGGGGACATTTATCTCAAGC 84  
QY 2010 TCAGGGCCCATCCCTTCTCTGACGCTGTGACCCAAATTTAGTCCAGAAATAACTGAGA 2069  
Db 83 TCAGGGCCCATCCCTTCTCTGACGCTGTGACCCAAATTTAGTCCAGAAATAACTGAGA 24  
QY 2070 AG 2071  
||  
Db 23 AG 22  
RESULT 11  
AY279380 1370 bp mRNA linear PRI 26-MAY-2003  
LOCUS Homo sapiens kallikrein 5 splice variant 1 (KLK5) mRNA, complete  
DEFINITION cds; alternatively spliced.  
ACCESSION AY279380  
VERSION AY279380.1 GI:31075480  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1370)  
AUTHORS Kurlender,L., Yousef,G.M., White,N.M.A., Robb,J.-D., Borgono,C.A.  
and Diamandis,E.P.  
TITLE Identification of splice variants for the human kallikrein gene 5  
(KLK5)  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1370)  
AUTHORS Kurlender,L., Yousef,G.M., White,N.M.A., Robb,J.-D., Borgono,C.A.  
and Diamandis,E.P.  
TITLE Direct Submission  
JOURNAL Submitted (19-APR-2003) Pathology and Laboratory Medicine, Mount  
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,  
Canada  
FEATURES  
source Location/Qualifiers  
1..1370  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.4"  
1..1370  
/gene="KLK5"  
1..127  
/gene="KLK5"  
128..1009  
/gene="KLK5"  
/note="serine protease; hK5; alternatively spliced;  
similar to the product of Genbank Accession Number  
AF135028"  
/codon\_start=1  
/product="kallikrein 5 splice variant 1"  
/protein\_id="AAP42275.1"  
/db\_xref="GI:31075481"  
/translation="MATARPMMVVCALITALLGVTEHVLANNVSCDHSFNPVS  
GSMQDVGAGGEDARSDSSRLINSDCDMHQPNQALILRPNQLYCGAVLVHQQW

LTAAHCKRKYFVRLGHYSLSPYESQOMFOGVKSIPHPGYSHPGSHNDMLIKLN  
RRIRPTKDVPRINSSHPSAGTKCLVSMGTGKSPVHPFKVLOCCLNLSYSOKRCE  
DAYPRIDDTMFECAGKXGRDSCQSDSGPVVNCNSLGLVSWBDYPCARRNRPQVYT  
NCKRTKNIQETIQANS"

3'UTR  
polyA-signal  
/gene="RLK5"  
/gene="RLK5"  
1010..1370  
1319..1324  
/gene="RLK5"

BASE COUNT 323 a 424 c 352 g 271 t  
ORIGIN

Query Match 23.3%; Score 482; DB 9; Length 1370;  
Best Local Similarity 100.0%; Pred. No. 8.9e-111;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGCGATTCGCGGGGGGCGCTGTGCTGCATGCTCCCTGCAGGAGCTGCTGCTGG 1649  
DB 851 CAGGCGATTCGCGGGGGGCGCTGTGCTGCATGCTCCCTGCAGGAGCTGCTGCTGG 910  
QY 1650 GAGAGTTACCTTGTGCCGGCCCAACAGACCGGGGTGTCTACAGCAACCTGTGCAAGTTG 1709  
DB 911 GAGAGTTACCTTGTGCCGGCCCAACAGACCGGGGTGTCTACAGCAACCTGTGCAAGTTG 970  
QY 1710 ACCAAGTGGATCCAGGAAACCAATCCAGGCAATCCTGAGTCATCCAGAGACTGACACA 1769  
DB 971 ACCAAGTGGATCCAGGAAACCAATCCAGGCAATCCTGAGTCATCCAGAGACTGACACA 1030  
QY 1770 CCGGATCCCGACCTGCTGCAGGAGACAGCCCTGACATCTCTTTCAGACCTCTATCTTC 1829  
DB 1031 CCGGATCCCGACCTGCTGCAGGAGACAGCCCTGACATCTCTTTCAGACCTCTATCTTC 1090  
QY 1830 CCAGAGATTTAGAGATGTTATCTCTCCAGCCCTGACCCCATGTCTCTGAGTCTGAGTCAAG 1889  
DB 1091 CCAGAGATTTAGAGATGTTATCTCTCCAGCCCTGACCCCATGTCTCTGAGTCTGAGTCAAG 1150  
QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGCTGTCTCTTCTAGTTGAACCTTGGGAACAATTTC 1949  
DB 1151 GTCTGCTTCCCGACATTTGGGCTGACCGCTGTCTCTTCTAGTTGAACCTTGGGAACAATTTC 1210  
QY 1950 CAAAAGTGTCCAGGAGGGGGGGTGGTGTCTCATCTCCCGGGGCACTTTCATCTCCCAAGC 2009  
DB 1211 CAAAAGTGTCCAGGAGGGGGGGTGGTGTCTCATCTCCCGGGGCACTTTCATCTCCCAAGC 1270  
QY 2010 TCAGGCGCCATCCCTTCTCTGAGCTGTGACCCCAATTAGTCCAGAAATAAAGTGAAGA 2069  
DB 1271 TCAGGCGCCATCCCTTCTCTGAGCTGTGACCCCAATTAGTCCAGAAATAAAGTGAAGA 1330  
QY 2070 AG 2071  
DB 1331 AG 1332

## RESULT 12

BD107879 1381 bp DNA linear PART 18-SEP-2002  
LOCUS BD107879 36 human secreted proteins.  
DEFINITION BD107879  
ACCESSION BD107879.1 GI:23202697  
VERSION  
KEYWORDS JP 2002500035-A/50.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1381)  
Ruben,S.M., Soppet,D.R., Ebner,R., Lafleur,D.W., Ni,J.,  
Brewer,L.A., Olsen,H.S., Duan,R.D. and Rosen,C.A.

36 human secreted proteins  
Patent: JP 2002500035-A 50 08-JAN-2002;  
HUMAN GENOME SCIENCES INC

OS Homo sapiens (human)  
PN JP 2002500035-A/50  
PD 08-JAN-2002  
PF 06-JAN-1999 JP 2000527554

COMMENT  
JP 2002500035-A/50  
PD 08-JAN-2002  
PF 06-JAN-1999 JP 2000527554

PR 07-JAN-1998 US 60/070657, 07-JAN-1998 US 60/070658 PR  
07-JAN-1998 US 60/070692, 07-JAN-1998 US 60/070704 PI STEVEN  
M RUBEN, DANIEL R SOPPET, REINHARD EBNER, DAVID W LAFLEUR, PI JIAN  
NI, LAURIE A BREWER, HENRIK S OLSEN, ROSANNE D DUAN, CRAIG A ROSEN PC  
PI 12N15/09, A61K31/711, A61K38/00, A61K39/393, A61K39/395, A61K48/00,  
PC A61P7/00, A61P11/06, A61P19/02, A61P29/00, A61P31/18, A61P35/00, PC  
A61P35/02,  
PC A61P37/00, C07K14/435, C07K16/18, C12N1/15, C12N1/21, C12N5/10, PC  
C12N15/00,  
PC A61K37/02, C12N5/00  
CC 36 human secreted proteins  
FH Key Location/Qualifiers  
FT source 1.1381  
FT /organism="Homo sapiens (human)".  
FEATURES  
source Location/Qualifiers  
1.1381  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 329 a 425 c 353 g 274 t  
ORIGIN

Query Match 23.3%; Score 482; DB 6; Length 1381;  
Best Local Similarity 100.0%; Pred. No. 8.9e-111;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGCGATTCGCGGGGGGCGCTGTGCTGCATGCTCCCTGCAGGAGCTGCTGCTGG 1649  
DB 852 CAGGCGATTCGCGGGGGGCGCTGTGCTGCATGCTCCCTGCAGGAGCTGCTGCTGG 911  
QY 1650 GAGAGTTACCTTGTGCCGGCCCAACAGACCGGGGTGTCTACAGCAACCTGTGCAAGTTG 1709  
DB 912 GAGAGTTACCTTGTGCCGGCCCAACAGACCGGGGTGTCTACAGCAACCTGTGCAAGTTG 971  
QY 1710 ACCAAGTGGATCCAGGAAACCAATCCAGGCAATCCTGAGTCATCCAGAGACTGACACA 1769  
DB 972 ACCAAGTGGATCCAGGAAACCAATCCAGGCAATCCTGAGTCATCCAGAGACTGACACA 1031  
QY 1770 CCGGATCCCGACCTGCTGCAGGAGACAGCCCTGACATCTCTTTCAGACCTCTATCTTC 1829  
DB 1032 CCGGATCCCGACCTGCTGCAGGAGACAGCCCTGACATCTCTTTCAGACCTCTATCTTC 1091  
QY 1830 CCAGAGATTTAGAGATGTTATCTCTCCAGCCCTGACCCCATGTCTCTGAGTCTGAGTCAAG 1889  
DB 1092 CCAGAGATTTAGAGATGTTATCTCTCCAGCCCTGACCCCATGTCTCTGAGTCTGAGTCAAG 1151  
QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGCTGTCTCTTCTAGTTGAACCTTGGGAACAATTTC 1949  
DB 1152 GTCTGCTTCCCGACATTTGGGCTGACCGCTGTCTCTTCTAGTTGAACCTTGGGAACAATTTC 1211  
QY 1950 CAAAAGTGTCCAGGAGGGGGGGTGGTGTCTCATCTCCCGGGGCACTTTCATCTCCCAAGC 2009  
DB 1212 CAAAAGTGTCCAGGAGGGGGGGTGGTGTCTCATCTCCCGGGGCACTTTCATCTCCCAAGC 1271  
QY 2010 TCAGGCGCCATCCCTTCTCTGAGCTGTGACCCCAATTAGTCCAGAAATAAAGTGAAGA 2069  
DB 1272 TCAGGCGCCATCCCTTCTCTGAGCTGTGACCCCAATTAGTCCAGAAATAAAGTGAAGA 1331  
QY 2070 AG 2071  
DB 1332 AG 1333

## RESULT 13

AF168768 1387 bp mRNA linear PRI 16-APR-2002  
LOCUS AF168768  
DEFINITION Homo sapiens stratum corneum tryptic enzyme (SCTE) mRNA, complete  
cds

ACCESSION AF168768  
VERSION AF168768.2 GI:20153423



polyA-signal 1387..1392  
/gene="KLK5"  
BASE COUNT 332 a 443 c 374 g 289 t  
ORIGIN

Query Match 23.3%; Score 482; DB 9; Length 1438;  
Best Local Similarity 100.0%; Pred. No. 9e-111;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1590 CAGGATGTTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGAGACTGTCTCTGG 1649
DB 919 CAGGATGTTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGAGACTGTCTCTGG 978
QY 1650 GGAGATTACCTCTTGGCCCGGCGCCACAGACGCGGGTGTCTACAGAACTCTGCAAGTTC 1709
DB 979 GGAATTAACCTTGTGCGCCGCGCCACAGACGCGGGTGTCTACAGAACTCTGCAAGTTC 1038
QY 1710 ACCAAGTGTATCAGAGAAACCATCCAGGCACTCTGAGTCAATCCAGAGACTGACACA 1769
DB 1039 ACCAAGTGTATCAGAGAAACCATCCAGGCACTCTGAGTCAATCCAGAGACTGACACA 1098
QY 1770 CCGGATCCCGACCTGCTGCAAGGAGACGCGCTGACACTCTTCAAGACCTCTCATCTTC 1829
DB 1099 CCGGATCCCGACCTGCTGCAAGGAGACGCGCTGACACTCTTCAAGACCTCTCATCTTC 1138
QY 1830 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCGCTGACCCCATGTCCTCTGAGTCAAG 1889
DB 1159 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCGCTGACCCCATGTCCTCTGAGTCAAG 1218
QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAACAATTTT 1949
DB 1219 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAACAATTTT 1278
QY 1950 CAAAGCTGTCCAGGGCGGGGGTGGTGGTCAATCTCCCTGGGGCACTTCACTCTCAAGC 2009
DB 1279 CAAAGCTGTCCAGGGCGGGGGTGGTGGTCAATCTCCCTGGGGCACTTCACTCTCAAGC 1338
QY 2010 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTAGTCCAGAAATTAACGTGAGA 2069
DB 1339 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTAGTCCAGAAATTAACGTGAGA 1398
QY 2070 AG 2071
DB 1399 AG 1400
```

RESULT 15  
BD107880 1439 bp DNA linear PAT 18-SEP-2002  
LOCUS  
DEFINITION 36 human secreted proteins.  
ACCESSION BD107880  
VERSION BD107880.1 GI:23202698  
KEYWORDS JP 2002500035-A/51.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 1439)  
AUTHORS Ruben,S.M., Soppet,D.R., Edner,R., Lafleur,D.W., NI,J.,  
Brewer,L.A., Olsen,H.S., Duan,R.D. and Rosen,C.A.  
36 human secreted proteins  
Patent: JP 2002500035-A 51 08-JAN-2002;  
TITLE HUMAN GENOME SCIENCES INC  
JOURNAL OS Homo sapiens (human)  
COMMENT PN JP 2002500035-A/51  
PD 08-JAN-2002  
PE 06-JAN-1999 JP 2000527554  
PR 07-JAN-1998 US 60/070657, 07-JAN-1998 US 60/070658 PR  
M 07-JAN-1998 US 60/070692, 07-JAN-1998 US 60/070704 PI STEVEN  
NI M RUBEN,DANIEL R SOPPET,REINHARD EBNER,DAVID W LAFLEUR, PI JIAN  
PI LAURIE A BREWER,HENRIK S OLSEN,ROSANNE D DUAN,CRAIG A ROSEN PC  
C12N15/09,A61K31/711,A61K38/00,A61K39/395,A61K39/395,A61K48/00.

PC A61P5/00,  
PC A61P7/00,A61P11/06,A61P19/02,A61P29/00,A61P31/18,A61P35/00, PC  
A61P35/02,  
PC A61P37/00,C07K14/435,C07K16/18,C12N1/15,C12N1/21,C12N5/10, PC  
C12N15/00,  
PC A61K37/02,C12N5/00  
CC 36 human secreted proteins  
FH key Location/Qualifiers  
FT source 1..1439  
FT /organism="Homo sapiens (human)".

FEATURES  
source  
1..1439  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 325 a 446 c 355 g 313 t  
ORIGIN

Query Match 23.3%; Score 482; DB 6; Length 1439;  
Best Local Similarity 100.0%; Pred. No. 9e-111;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1590 CAGGATGTTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGAGACTGTCTCTGG 1649
DB 910 CAGGATGTTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGAGACTGTCTCTGG 969
QY 1650 GGAGATTACCTTGTGCGCCGCGCCACAGACGCGGGTGTCTACAGAACTCTGCAAGTTC 1709
DB 970 GGAATTAACCTTGTGCGCCGCGCCACAGACGCGGGTGTCTACAGAACTCTGCAAGTTC 1029
QY 1710 ACCAAGTGTATCAGAGAAACCATCCAGGCACTCTGAGTCAATCCAGAGACTGACACA 1769
DB 1030 ACCAAGTGTATCAGAGAAACCATCCAGGCACTCTGAGTCAATCCAGAGACTGACACA 1089
QY 1770 CCGGATCCCGACCTGCTGCAAGGAGACGCGCTGACACTCTTCAAGACCTCTCATCTTC 1829
DB 1090 CCGGATCCCGACCTGCTGCAAGGAGACGCGCTGACACTCTTCAAGACCTCTCATCTTC 1149
QY 1830 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCGCTGACCCCATGTCCTCTGAGTCAAG 1889
DB 1150 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCGCTGACCCCATGTCCTCTGAGTCAAG 1209
QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAACAATTTT 1949
DB 1210 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAACAATTTT 1269
QY 1950 CAAAGCTGTCCAGGGCGGGGGTGGTGGTCAATCTCCCTGGGGCACTTCACTCTCAAGC 2009
DB 1270 CAAAGCTGTCCAGGGCGGGGGTGGTGGTCAATCTCCCTGGGGCACTTCACTCTCAAGC 1329
QY 2010 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTAGTCCAGAAATTAACGTGAGA 2069
DB 1330 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTAGTCCAGAAATTAACGTGAGA 1389
QY 2070 AG 2071
DB 1390 AG 1391
```

Search completed: August 5, 2003, 09:57:09  
Job time : 7380.88 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 5, 2003, 05:12:55 ; Search time 129.698 seconds  
(without alignments)  
6806.329 Million cell updates/sec

Title: US-09-936-271B-13\_COPY\_1\_2000

Perfect score: 2000

Sequence: 1 gggccagagtgaggcaag.....gtgcgtctgcaccacatc 2000

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB-seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCRTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	444	22.2	246240	2	US-08-724-394A-20
C 2	444	22.2	246240	2	US-08-724-394A-21
C 3	444	22.2	246240	2	US-08-724-394A-22
C 4	380.4	19.0	246240	2	US-08-724-394A-20
C 5	380.4	19.0	246240	2	US-08-724-394A-21
C 6	380.4	19.0	246240	2	US-08-724-394A-22
C 7	366.4	18.3	70000	4	US-09-851-896-3
C 8	363.6	18.2	84495	4	US-09-797-906-3
C 9	362.8	18.1	14581	4	US-08-520-373D-4
C 10	362.8	18.1	22481	4	US-08-367-841A-43
C 11	362.8	18.1	22481	5	PCT-US95-07201-43
C 12	362.8	18.1	22484	4	US-09-875-223-2
C 13	333.6	17.7	5375	3	US-08-757-223-7
C 14	333.6	17.7	39982	4	US-09-820-924-3
C 15	343.6	17.2	116592	4	US-09-818-512-3
C 16	339.6	17.0	3805	4	US-09-108-006C-3
C 17	339.2	17.0	168575	4	US-09-426-290-1
C 18	337.2	16.9	29629	4	US-09-729-995-3
C 19	336.6	16.8	43950	4	US-09-735-934A-3
C 20	336.6	16.8	43950	4	US-10-060-332-3
C 21	333.6	16.7	174493	4	US-09-804-471A-3
C 22	332.8	16.6	18073	3	US-09-078-294-12
C 23	332.6	16.6	43950	4	US-09-735-934A-3
C 24	332.6	16.6	43950	4	US-10-060-332-3
C 25	331	16.6	19011	1	US-08-310-356-36
C 26	331	16.6	19557	5	PCT-US92-06300-1
C 27	330.8	16.5	70000	4	US-09-851-896-3

Sequence 10, Appl  
Sequence 15, Appl  
Sequence 18, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 10, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 18, Appl  
Sequence 4, Appl  
Sequence 43, Appl  
Sequence 43, Appl  
Sequence 2, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-724-394A-20/c  
; Sequence 20, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Kronmal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Tsuchihashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,394A  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitts, Renee A.  
; REGISTRATION NUMBER: 35,136  
; REFERENCE/DOCKET NUMBER: 017957-000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 246240 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: cdna  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..246240  
; OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-20















```

1  PRIOR APPLICATION DATA:
2  APPLICATION NUMBER: 07/952,796
3  FILING DATE: 24-SEP-1992
4  ATTORNEY/AGENT INFORMATION:
5  NAME: DOROTHY R. AUTH
6  REGISTRATION NUMBER: 36434
7  REFERENCE/DOCKET NUMBER: 20264
8  TELECOMMUNICATION INFORMATION:
9  TELEPHONE: (212) 758-4800
10 TELEFAX: (212) 751-6849
11 INFORMATION FOR SEQ ID NO: 43:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 22481 Base Pairs
14 TYPE: Nucleic Acid
15 STRANDEDNESS: Double
16 TOPOLOGY: Unknown
17 MOLECULE TYPE: Genomic DNA
18 FEATURE:
19 NAME/KEY: pl-147
20 LOCATION:
21 IDENTIFICATION METHOD:
22 OTHER INFORMATION: full length
23 OTHER INFORMATION: sequence for
24 US-08-367-841A-43

```

Query Match	18.1%;	Score 362.8;	DB 4;	Length 22481;
Best Local Similarity	76.3%;	Pred. No. 5.9e-65;		
Matches 528;	Conservative 0;	Mismatches 97;	Indels 67;	Gaps 4;

Qy	359	TTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTCCGCCAGGCTCGAGTGCAGTGGC	418
Db	14099	TTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTCCGCCAGGCTCGAGTACGGTGGC	14040
Qy	419	GGGATCTCGGCTCACTGCAAGCTCCGGCTTCCCGGGTTCAAGCCATTTCCTGCCTCAGCC	478
Db	14039	GTGATCTCGGCTCACTGCAAGCTCCGCTCCAGGTTCAAGCCATTTCCTGCCTCAGCC	13980
Qy	479	TCCCAAGTAGCTGGGACTACAGGCCGCCGCCACTAGGCCGGCTCAATTTTTTTGTATTTT	538
Db	13979	TCCCGAGTAGCTGGGACTACAGGCCGCCACCACCATGCTGGGTAA-TTTTTTGTATTTT	13921
Qy	539	TAGTAGAGACGGGTTTACCGTTTATGCCGGGATGGCTCGATCTCTGACCTCGTGAT	598
Db	13920	TAGTAGAGACGGGTTTACCGTTTATGACGAGGATGGTCTCGATCTCTGACCTCGTGAT	13861
Qy	599	CCGCCCGCTCGGCCCTCCCAAAGTGTGGGATTTACAGCGGTGAGCCACCGCGCCGGCCA	658
Db	13860	CCACCCGCTCGGCTTACCAAAGTGTGGGATTTACAGGCATGAGCCACCGCGCCGAGC--	13803
Qy	659	TGATCATCTTCTTGACTATGCTGATGTGAAGTACCTAAAGCCATCAGACTCTACCCCTT	718
Db	13802	-----CCATGACCCACTTTTTTAAAAACCATCA-----	13775
Qy	719	TAAATATGACGTTTGGGCCGAGGACCGGTGCCTGATGCTGTAAATTCACGACTTTGGAG	778
Db	13774	-----GCCGAGTGGCTCACACCTATAATCCACGACTTTGGAG	13736
Qy	779	GCAGAGGTGGTGAATCACTTCAGGCCAGGAGTTTTCAGACCAGCTGCCAACATGGTA	838
Db	13735	GCTGAGGTGGTGAATCACTTCAGGTTCAGGAGTTTCGAGACCATCTCGGCCAACACGCTGA	13676
Qy	839	AACTCTGTCTTTACTAAAAAATAAAAAAAAAAAAAATTCAGCGGGTCTCGTGGGGC	898
Db	13675	AAACCGTCTCT-----ACTAAAAATACAAAAAATTAACGGGGCATGGTGTGC	13627
Qy	899	ACACCTGTATCCAGCTATGCTGGAGGCTGAGGCACGAGATCACTTTGAACCTCGAGG	958
Db	13626	ATGCTGTATCCAGCTACTCAGGAGGCTGAGGCACGAGATTCCTGTTGAACCCAGGAGG	13567
Qy	959	CGAGGTTGCAGTGGGCCGAGATCACATCACCGCCCTCCAGCTGGGGCAGACAGACAAGA	1018
Db	13566	CAGAGTTGCAGTGCAGTGCAGATGATGCCACTGTACTCCAGCTGGGTGACAGAGTGA	13507

```

QY      1019  CTCTGTCTCAATAAATAATAACAACACGAA 1050
      | | | | | | | | | | | | | | | | | | | |
Db      13506  CCTGTGTCACAAAAATAAATAAGTAATAAATAA 13475

RESULT 11
PCT-US95-07201-43/G
: Sequence 43, Application PC/TUS9507201
: GENERAL INFORMATION:
: APPLICANT: Chader, Gerald J.; Becerra, Sofia
: APPLICANT: Patricia; Schwartz, Joan P.;
: APPLICANT: Taniwaki, Takayuki
: TITLE OF INVENTION: PIGMENT EPITHELIUM
: TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
: TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morgan & Finnegan, L.L.P.
: STREET: 345 Park Avenue
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10154
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy Disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORDPERFECT 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/07201
: FILING DATE: 06-JUN-1995
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/367,841
: FILING DATE: 30-DEC-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/257,963
: FILING DATE: 07-JUN-1994
: PRIOR APPLICATION DATA: 07/952,796
: APPLICATION NUMBER: 24-SEP-1992
: FILING DATE: 24-SEP-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: DOROTHY R. AUTH
: REGISTRATION NUMBER: 36434
: REFERENCE/DOCKET NUMBER: 20264126PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 758-6800
: TELEFAX: (212) 751-6849
: INFORMATION FOR SEQ ID NO: 43:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 22481 Base Pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Double
: TOPOLOGY: Unknown
: MOLECULE TYPE: Genomic DNA
: FEATURE:
: NAME/KEY: PI-147
: LOCATION:
: IDENTIFICATION METHOD:
: OTHER INFORMATION: full length genomic
: OTHER INFORMATION: sequence for PEDF plus flanking sequences
PCT-US95-07201-43

```

	Query Match	18.1%	Score 362.8;	DB 5;	Length 22481;
	Best Local Similarity	76.3%;	Pred. No. 5.9e-65;		
	Matches 528;	Conservative 0;	Mismatches 97;	Indels 67;	Gaps 4;
Qy	359	TTTTTTTTTTTTTTTTTTGAGCGGAGTCTCGCTCTGTCCGCCAGCGCTGGAGTCAGTGGC	418		
Db	14099	TTTTTTTTTTTTTTTTTTGAGCGGAGTCTCGCTCTGTCCGCCAGCGCTGGAGTACGGTGGC	14040		
Qy	419	GGGATCTCGGCTCACTGCAAGCTCGGCCCTCCGGGTTACCGCCCATTCCTTCCTCGCTCAGCC	478		







**THIS PAGE BLANK (USPTO)**





J. Biol. Chem. 274 (53), 37511-37516 (1999)  
 20076408  
 MEDLINE  
 PUBMED  
 10608802  
 2 (bases 1 to 11570)  
 Diamandis,E.P., Yousef,G.M., Luo,L.Y., Magklara,A. and Obiezu,C.V.  
 The new human kallikrein gene family: implications in  
 carcinogenesis  
 Trends Endocrinol. Metab. 11 (2), 54-60 (2000)  
 21121728  
 MEDLINE  
 PUBMED  
 10675891  
 3 (bases 1 to 11570)  
 Yousef,G.M., Luo,L.Y. and Diamandis,E.P.  
 Direct Submission  
 Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount  
 Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,  
 Canada

FEATURES  
 source  
 1. .11570  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /map="19q13.3-q13.4"  
 join(2101..2131,2210..2293,4762..5023,5763..6019,  
 6105..6238,11092..11570)  
 /product="kallikrein-like protein 2 KLK-L2"  
 join(2221..2293,4762..5023,5763..6019,6105..6238,  
 11092..11247)  
 /codon\_start=1  
 /product="kallikrein-like protein 2 KLK-L2"  
 /protein\_id="AAD26429.1"  
 /db\_xref="GI:4589283"  
 /translation="MATARPWMVLCALITALLGVTEHVLANNVSCDHPNSTVPS  
 GSNQDLGAGEDARSDSSRIINGSDCMHPQWQAALLRPNQLKCAVLPVPOW  
 LLTAHCKKRVFRLYRLSLSPVIESGQMFQVRSIPHPHSHQSGHSLNDLMLKLN  
 RRTPTKDVPIINVSSPCPSAGTKCLVSGWGTTSKSPQVHPFKVLQCLINISVLSQKRC  
 DAYPRQIDDTMFCAGDKRAGRDSCQDGGPVCVNGSLQGLVSWGIDYPCARPNRPGVYT  
 NLCKFTKIOWETQANS"  
 BASE COUNT 3034 a. 2562 c 3325 g 2649 t  
 ORIGIN

Query Match 100.0%; Score 2000; DB 9; Length 11570;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGCCAGAGTGAAGCAAGAGAGAGTTCAGAGCTCCCTCTCAAGTGGCTTGATC 60  
 Db 1 GGGCCAGAGTGAAGCAAGAGAGAGTTCAGAGCTCCCTCTCAAGTGGCTTGATC 60  
 Qy 61 TCCCTGCCTAAATTCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
 Db 61 TCCCTGCCTAAATTCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
 Qy 121 AAGAAAGAGAGAGAGAGAGAGACAGAAATACAACTACAGAAACACAGAGAGACAC 180  
 Db 121 AAGAAAGAGAGAGAGAGAGAGACAGAAATACAACTACAGAAACACAGAGAGACAC 180  
 Qy 181 ACAGAGAGCTGGGACACAGGACACAGAGTACAGAGAGAGAGAGAGAGAGAGAG 240  
 Db 181 ACAGAGAGCTGGGACACAGGACACAGAGTACAGAGAGAGAGAGAGAGAGAGAG 240  
 Qy 241 ACACAAATGGAGACACAGAGTGTAAAGAAAGAGAGATTAACAGAGTCCAGATACACGC 300  
 Db 241 ACACAAATGGAGACACAGAGTGTAAAGAAAGAGAGATTAACAGAGTCCAGATACACGC 300  
 Qy 301 AAAGGGCAGAGACAGAGTTTACGGTGGTGTCTATGATCATCTCTTTTTTTTTTT 360  
 Db 301 AAAGGGCAGAGACAGAGTTTACGGTGGTGTCTATGATCATCTCTTTTTTTTTTT 360  
 Qy 361 TTTTTTTTTTTTTTTGAGCGGAGCTCTCGCTCTGCGCCAGGCTCGAGTGCAGTGGCGG 420  
 Db 361 TTTTTTTTTTTTTTTGAGCGGAGCTCTCGCTCTGCGCCAGGCTCGAGTGCAGTGGCGG 420

Qy 421 GATCTCGGCTCACTGCAAGCTCCGCCCTCCGGGGTTACGGCATTTCTCTGCTCAGCCTC 480  
 Db 421 GATCTCGGCTCACTGCAAGCTCCGCCCTCCGGGGTTACGGCATTTCTCTGCTCAGCCTC 480  
 Qy 481 CCAAGTAGCTGGGACTACAGCGCCGCCACTACGCCGGGTAATTTTTTTTGTATTTT 540  
 Db 481 CCAAGTAGCTGGGACTACAGCGCCGCCACTACGCCGGGTAATTTTTTTTGTATTTT 540  
 Qy 541 GTAGAGAGGGGTTTACCGCTTTAGCGGGATGGCTCGATCTCCTGACCTCTGATCC 600  
 Db 541 GTAGAGAGGGGTTTACCGCTTTAGCGGGATGGCTCGATCTCCTGACCTCTGATCC 600  
 Qy 601 GCCCGCTCGCCCTCCCAAGTGTGAGGCTGAGCGCTGAGCCACCGCCGCCCATG 660  
 Db 601 GCCCGCTCGCCCTCCCAAGTGTGAGGCTGAGCGCTGAGCCACCGCCGCCCATG 660  
 Qy 661 ATCATCTTGTGACTATGCTGATGACAGTACTAAAGCATCAGACTCTACCCTTA 720  
 Db 661 ATCATCTTGTGACTATGCTGATGACAGTACTAAAGCATCAGACTCTACCCTTA 720  
 Qy 721 AATATGAGTTTGGCCAGGACCGTGGCTCATGCTGTAATTCAGCAGCTTTGGAGGC 780  
 Db 721 AATATGAGTTTGGCCAGGACCGTGGCTCATGCTGTAATTCAGCAGCTTTGGAGGC 780  
 Qy 781 AGAGTGGTGAATCACTTGAAGGACGAGGTTTGAAGCCAGCTTGGCCAACTGTTGAAA 840  
 Db 781 AGAGTGGTGAATCACTTGAAGGACGAGGTTTGAAGCCAGCTTGGCCAACTGTTGAAA 840  
 Qy 841 CTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 900  
 Db 841 CTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 900  
 Qy 901 ACCTGTAATCCAGCTATGCTGGAGCTGAGGACGAGAGTCACTTCAACCTCGAGGCG 960  
 Db 901 ACCTGTAATCCAGCTATGCTGGAGCTGAGGACGAGAGTCACTTCAACCTCGAGGCG 960  
 Qy 961 GAGGTTCCAGTGGGCCGAGATACATCACCGCCCTCCAGCTGGGCGACAGAGCAAGCT 1020  
 Db 961 GAGGTTCCAGTGGGCCGAGATACATCACCGCCCTCCAGCTGGGCGACAGAGCAAGCT 1020  
 Qy 1021 CTGCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1080  
 Db 1021 CTGCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1080  
 Qy 1081 AAAAAAATAATGCTCAACAAATAGAGCAGAGTGAATAAAGAAATAAATAAATAAATAA 1140  
 Db 1081 AAAAAAATAATGCTCAACAAATAGAGCAGAGTGAATAAAGAAATAAATAAATAAATAA 1140  
 Qy 1141 AGAACTTAAGTATATTGACAAATCAATTCAGAACCTTTAAAAAAGAAAGAACACAGA 1200  
 Db 1141 AGAACTTAAGTATATTGACAAATCAATTCAGAACCTTTAAAAAAGAAAGAACACAGA 1200  
 Qy 1201 GGCATAGAAAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
 Db 1201 GGCATAGAAAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
 Qy 1261 ACAAGGCTCTTAAGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
 Db 1261 ACAAGGCTCTTAAGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
 Qy 1321 AAAAAGACAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
 Db 1321 AAAAAGACAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
 Qy 1381 GAGAGAGGAGGAGAGAGACAGAGATATTGAGAGAGACTCAAAAGATAGCCGAGGGA 1440  
 Db 1381 GAGAGAGGAGGAGAGAGACAGAGATATTGAGAGAGACTCAAAAGATAGCCGAGGGA 1440  
 Qy 1441 GAACACAGAGAGATGAAGAAGACTCTGAGAAAAAACACAGAGACAAAGATGGAAAGAG 1500  
 Db 1441 GAACACAGAGAGATGAAGAAGACTCTGAGAAAAAACACAGAGACAAAGATGGAAAGAG 1500  
 Qy 1501 AGTATCGAGGGTGAACAGACAGTGGTGAATAGCAAAATGACAGAAAGAAAGCAAGCAA 1560

```
Db 1501 AGTATCGAGGGTGAACAGACAGTGTGGTAATGAGCAAAATGACAGAAAGCAAGCAAA 1560
QY 1561 TCCAGCGCCCAAGATAGTACCCAGAGTTGTTGAGAACCCAGATCCTTAAGGCTGGGGG 1620
Db 1561 TCCAGCGCCCAAGATAGTACCCAGAGTTGTTGAGAACCCAGATCCTTAAGGCTGGGGG 1620
QY 1621 AGGCAGGGAAGGGGCTGGCTGGCTGCCGAGACCCCTCCCATTTCTCCGGGCCAGGGAG 1680
Db 1621 AGGCAGGGAAGGGGCTGGCTGGCTGCCGAGACCCCTCCCATTTCTCCGGGCCAGGGAG 1680
QY 1681 GTAGGAGTGACATCCCGGACTGGGTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCA 1740
Db 1681 GTAGGAGTGACATCCCGGACTGGGTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCA 1740
QY 1741 GGAGGAGCTATTGCTTAAGGCCCGATAGGACACCTCATTTGCCCGGGATGTGCCCCAGGGAG 1800
Db 1741 GGAGGAGCTATTGCTTAAGGCCCGATAGGACACCTCATTTGCCCGGGATGTGCCCCAGGGAG 1800
QY 1801 CAGTGGGTGTTTAACTCAGGCCCGGTGCCAGAGCCCGAGGAGGAGGAGTGGCCAGGA 1860
Db 1801 CAGTGGGTGTTTAACTCAGGCCCGGTGCCAGAGCCCGAGGAGGAGGAGTGGCCAGGA 1860
QY 1861 AGGCACAGCCCTGAGAAAGTCTCGGCTGAGCTGGGAGCAAAATCCCCACCCCTACCTGG 1920
Db 1861 AGGCACAGCCCTGAGAAAGTCTCGGCTGAGCTGGGAGCAAAATCCCCACCCCTACCTGG 1920
QY 1921 GGCACAGGCAAGTGAACCTGGTGGGTGGCTCAGCAGGAGGAGGAGGAGGTGTCT 1980
Db 1921 GGCACAGGCAAGTGAACCTGGTGGGTGGCTCAGCAGGAGGAGGAGGAGGTGTCT 1980
QY 1981 GTGGCTCTGTCACCCACATC 2000
Db 1981 GTGGCTCTGTCACCCACATC 2000
```

```
RESULT 2
AC011483/c
LOCUS 107487 bp DNA linear PRI 01-JUL-2002
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.
AC011483
VERSION AC011483.7 GI:21637461
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 107487)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished
2 (bases 1 to 107487)
DOE Joint Genome Institute.
Direct Submission
Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 107487)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jul 1, 2002 this sequence version replaced gi:14971176.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.6.
NOTE: Shatter libraries failed to resolve dinucleotide repeat.
Unsure number of repeat copies 64998-65494. Forced join 65015.
Location/Qualifiers
1. .107487
/organism="Homo sapiens"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-147C22"
64998..65494
/note="NOTE: Shatter libraries failed to resolve
dinucleotide repeat. Unsure number of repeat copies
64998-65494. Forced join 65015."
BASE COUNT 26044 a 27304 c 25728 g 28411 t
ORIGIN
Query Match 99.5%; Score 1989; DB 9; Length 107487;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2000; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GGGCCCCAGAGTGAAGGCAAGAGAGTTGAGAGCTCCCTCTGCAAAAGTGGCTTGAATC 60
Db 60220 GGGCCCCAGAGTGAAGGCAAGAGAGTTGAGAGCTCCCTCTGCAAAAGTGGCTTGAATC 60161
QY 61 TCCCTCTGCCCTAAATGTCAGGGAGAGGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 119
Db 60160 TCCCTCTGCCCTAAATGTCAGGGAGAGGGCAGAGAGAGAGAGAGAGAGAGAGAGAG 60101
QY 120 GAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 179
Db 60100 GAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60041
QY 180 CACAGAGAGCCCTGGGACACAGGAGACACAGAGAGTTCAGAGAGAGAGAGAGAGAGAG 239
Db 60040 CACAGAGAGCCCTGGGACACAGGAGACACAGAGAGTTCAGAGAGAGAGAGAGAGAG 59981
QY 240 GACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGATTAAACAGAGTCCCAGATAC 299
Db 59980 GACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGATTAAACAGAGTCCCAGATAC 59921
QY 300 CAAAGGGCAGAGACACAGTTCAGGGTGGTGTCTATGATCATCTCTCTTTTCTTTT 359
Db 59920 CAAAGGGCAGAGACACAGTTCAGGGTGGTGTCTATGATCATCTCTCTTTTCTTTT 59861
QY 360 TTTTCTTTTCTTTTGTGAGACGAGTCTCGTCTGTGCGCCAGGCTGGAGTGGAGTGGCG 419
Db 59860 TTTTCTTTTCTTTTGTGAGACGAGTCTCGTCTGTGCGCCAGGCTGGAGTGGAGTGGCG 59801
QY 420 GGATCTCGGCTCACTGCAAGCTCCGCTCCCGGTTTCAGCCACTTCTCTGCTCAGCCT 479
Db 59800 GGATCTCGGCTCACTGCAAGCTCCGCTCCCGGTTTCAGCCACTTCTCTGCTCAGCCT 59741
QY 480 CCCAAGTAGCTGGGACTACAGGCGCCCGCCACTACCGCGGCTAAATTTTGTATTTT 539
Db 59740 CCCAAGTAGCTGGGACTACAGGCGCCCGCCACTACCGCGGCTAAATTTTGTATTTT 59681
QY 540 AGTAGAGCGGGTTTACCGGTTTACCGGGATGCGCTCGATCTCCTGACCTCGTATC 599
Db 59680 AGTAGAGCGGGTTTACCGGTTTACCGGGATGCGCTCGATCTCCTGACCTCGTATC 59621
QY 600 CCGCCGCTCGGCTCCCAAGTGTGAGATACAGGCGTGGAGCCGACGCGCCGCCCAT 659
Db 59620 CCGCCGCTCGGCTCCCAAGTGTGAGATACAGGCGTGGAGCCGACGCGCCGCCCAT 59561
QY 660 GATCATCTTCTTGTACTATGCTGATGTGACAAGTACCTAAAGCCATCAGACTCTAC 719
Db 59560 GATCATCTTCTTGTACTATGCTGATGTGACAAGTACCTAAAGCCATCAGACTCTAC 59501
QY 720 AAATATGAGTTTGGGCCAGGACCGTGGCTCATGCTGCTAAATTCAGACACTTTGGGAG 779
Db 59500 AAATATGAGTTTGGGCCAGGACCGTGGCTCATGCTGCTAAATTCAGACACTTTGGGAG 59441
QY 780 CAGAGTGGTGAATCAGTTCAGGCGAGGTTTGAACAGCAGCTTCGCAACATGTGAA 839
Db 59440 CAGAGTGGTGAATCAGTTCAGGCGAGGTTTGAACAGCAGCTTCGCAACATGTGAA 59381
QY 840 ACTCTGCTTTTACTAAAAAATAAAAAAAAAAAAAAAAAAAAAATCAGCCGGGTGCTCGTGGGCA 899
Db 59380 ACTCTGCTTTTACTAAAAAATAAAAAAAAAAAAAAAAAAAAAATCAGCCGGGTGCTCGTGGGCA 899
```

Db 59380 ACTGTGCTTTACTATAAAAAAAAAAAAAAAAAAAAAAAAAAATACAGCCGGGTGTCTGTGGGCA 59321  
QY 900 CACCTGTAAATCCAGCTATCTGAGCTGAGGACGAGAGTCACTTGAACCTGGAGGC 959  
Db 59320 CACCTGTAAATCCAGCTATCTGAGGCTGAGGACGAGAGTCACTTGAACCTGGAGGC 59261  
QY 960 GGAGTTGTCAGTGGCGGAGATCATACCGCCCTCCAGCCTGGCGGAGAGCAAGAC 1019  
Db 59260 GGAGTTGTCAGTGGCGGAGATCATACCGCCCTCCAGCCTGGCGGAGAGCAAGAC 59201  
QY 1020 TCTCTCAATAAATAAATAAACAAGCAAGCAGTGTGTGTACCTTGTATATC 1079  
Db 59200 TCTCTCAATAAATAAATAAACAAGCAAGCAGTGTGTGTACCTTGTATATC 59141  
QY 1080 TAAAAAATAATGCTGTCAACATAGCAGAGTGAATATAAGGAAATAAATGCGCC 1139  
Db 59140 TAAAAAATAATGCTGTCAACATAGCAGAGTGAATATAAGGAAATAAATGCGCC 59081  
QY 1140 AAGAACTCTAAGTATATTTGACAAATCATTCAGAACCCTTTAAAAAAGAAGATCAG 1199  
Db 59080 AAGAACTCTAAGTATATTTGACAAATCATTCAGAACCCTTTAAAAAAGAAGATCAG 59021  
QY 1200 AGCATAGAAAGACAGGAGGAAACAGGAGACAGAAACACCTGTGGCCCAAGGAGACAA 1259  
Db 59020 AGCATAGAAAGACAGGAGGAAACAGGAGACAGAAACACCTGTGGCCCAAGGAGACAA 58961  
QY 1260 AACAGGCTCTTAAGACAGACAGGAGGAGAGAGAGTGTAGTGAGACAGACAGACA 1319  
Db 58960 AACAGGCTCTTAAGACAGACAGGAGGAGAGAGAGTGTAGTGAGACAGACAGACA 58901  
QY 1320 GAAAAACAGAGAGAGAGACAGACAGACAGAGAGCGGAGAGGGATAGAAAG 1379  
Db 58900 GAAAAACAGAGAGAGAGACAGACAGACAGAGAGCGGAGAGGGATAGAAAG 58841  
QY 1380 AGCAGAGGGGTGAGAGACAGACAGATATTGAGAGAGTCTAGAAGATAGCCGAGG 1439  
Db 58840 AGCAGAGGGGTGAGAGACAGACAGATATTGAGAGAGTCTAGAAGATAGCCGAGG 58781  
QY 1440 AGAACACAGAGAGATGAAGAGACTCTGAGAAAAAACAGACAGACAAAGATGAAAG 1499  
Db 58780 AGAACACAGAGATGAAGAGACTCTGAGAAAAAACAGACAGACAAAGATGAAAG 58721  
QY 1500 GAGTATCGAGGTGAACAGACAGTGTGTGAATAGCAAAATGACAGAGAAAGCAAGCA 1559  
Db 58720 GAGTATCGAGGTGAACAGACAGTGTGTGAATAGCAAAATGACAGAGAAAGCAAGCA 58661  
QY 1560 ATCCAGGCGCAACAATAGTGACCCAGAGTTGGTGGAAGCCAGATCCTTAAGGCTGGG 1619  
Db 58660 ATCCAGGCGCAACAATAGTGACCCAGAGTTGGTGGAAGCCAGATCCTTAAGGCTGGG 58601  
QY 1620 GAGCAGGGAAGGGGCTGGCTTCCGAGAGACCCCTCCCAATTCCTCGGGCCAGGGA 1679  
Db 58600 GAGCAGGGAAGGGGCTGGCTTCCGAGAGACCCCTCCCAATTCCTCGGGCCAGGGA 58541  
QY 1680 GGTAGGAGTACATTCGGACTGGGTGGGGGTGCTCTGGGGGTGAGATAGGGGAGC 1739  
Db 58540 GGTAGGAGTACATTCGGACTGGGTGGGGGTGCTCTGGGGGTGAGATAGGGGAGC 58481  
QY 1740 AGGAGGAGCTATTGCTAAGGCGCCATAGGCACTTATGTCGGCGGAATGTGCCCGAGGGA 1799  
Db 58480 AGGAGGAGCTATTGCTAAGGCGCCATAGGCACTTATGTCGGCGGAATGTGCCCGAGGGA 58421  
QY 1800 GCAGTGGTGGTTAATACTACGCGCCGCTGCCAGAGCCCAAGGAGGAGGAGTGGCCAGG 1859  
Db 58420 GCAGTGGTGGTTAATACTACGCGCCGCTGCCAGAGCCCAAGGAGGAGGAGTGGCCAGG 58361  
QY 1860 AAGCAGAGGCTGAGAAAGTCTGCGCTGAGCTGGGAGCAATCCGCCACCCCTACCTG 1919  
Db 58360 AAGCAGAGGCTGAGAAAGTCTGCGCTGAGCTGGGAGCAATCCGCCACCCCTACCTG 58301  
QY 1920 GGGGACAGGCGCAAGTGTGAGGCTGTGAGGCTGTGAGGAGGAGGAGAGGAGTGTCT 1979  
Db 58300 GGGGACAGGCGCAAGTGTGAGGCTGTGAGGCTGTGAGGAGGAGGAGGAGGAGTGTCT 58241

QY 1980 TGTGCGTCTTGACCCACATC 2000  
Db 58240 TGTGCGTCTTGACCCACATC 58220

RESULT 3  
AF243527/c 230000 bp DNA linear PRI 21-NOV-2000  
LOCUS Homo sapiens serine protease gene cluster, complete sequence.  
DEFINITION AF243527  
ACCESSION AF243527  
VERSION AF243527.1 GI:11244757  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 230000)  
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuall, J.,  
Moss, P., Paepel, B. and Wang, K.  
TITLE Sequencing and expression analysis of the serine protease gene  
cluster located in chromosome 19q13 region  
JOURNAL Gene 257 (1), 119-130 (2000)  
MEDLINE 20510030  
PUBMED 11054574  
REFERENCE 2 (bases 1 to 230000)  
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuall, J.,  
Moss, P., Paepel, B. and Wang, K.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,  
Bothell, WA 98021, USA  
FEATURES  
source  
1. 230000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13"  
complement(<6963..>11517)  
/gene="KLK1"  
complement(join(<6963..7118,7668..7804,7923..8212,  
9482..9641,11472..>11517))  
/gene="KLK1"  
/product="renal kallikrein"  
complement(join(6963..7118,7668..7804,7923..8212,  
9482..9641,11472..>11517))  
/gene="KLK1"  
/note="serine protease"  
/codon\_start=1  
/product="renal kallikrein"  
/protein\_id="AAG33353.1"  
/db\_xref="GI:11244758"  
/translation="MWFLVLCALSLGSGTAAPDQSRIVGGWECEHSPQWQAALVH  
FSFQCGILVHROWLTAHAICSDNYQLHGRHNLFDDETAQFVHVSFPHPGFN  
MSLNEFRDADEYSHDLMLRLTEPADTIDRAVKVVELPEPEVGSCTCLASGWS  
IPEFNFSFDLQCDLKLILPNDCKRAHKVQYDFMLCVGHLEGGKDTGVDSGGPL  
MCDVGLQGVSWGVPGCTPNKPSVAVRYSYVKWIEDTIAENS"  
complement(join(<13552..13704,14377..14530,14678..>14917,  
15416..15569,19204..>19246))  
/product="ACO protease"  
complement(join(13552..13704,14377..14530,14678..>14917,  
15416..15569,19204..>19246))  
/note="serine protease"  
/codon\_start=1  
/product="ACO protease"  
/protein\_id="AAG33354.1"  
/db\_xref="GI:11244759"  
/translation="MWLLTFLSLASTAQAQDKLLEGDECAPHSQWQVALYERGR  
FNCGASLIQHWLVAACQSRFRVRLGHNLRKRDGPOLRTTSTVIRPHRYEARS  
HRNDIMLLRVLQPARLPQVRPAVLTRCPHVGACVSGWGLVPLSPVSLPTLHC  
ANISIIIDTCDKSYPGRLNTMVCAGABRGABSCGDSGGPLVCGGILQGVISWGD  
VPDNTTKPGVYTKVCHYLEWIRETKRN"  
<42595..>47769  
mRNA  
CDS  
gene

[illegible]



JOURNAL  
REFERENCE  
AUTHORS

Unpublished  
2 (bases 1 to 217346)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,  
Campiolo,A., Castile,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,P., Ferreira,P., FitzHugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Headford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,  
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,  
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Menga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tessaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 217346)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,  
Boukhgalter,B., Brown,A., Burkett,G., Campiolo,A., Castile,A.,  
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,  
DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,  
Graham,L., Grand-Pierre,N., Hagos,B., Headford,A., Horton,L.,  
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,  
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,  
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Menga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tessaye,S., Theodore,J.,  
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.

TITLE  
JOURNAL

Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Nov 15, 2000 this sequence version replaced gi:11136831.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

## COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L9166  
Center clone name: 795\_B\_6  
----- Summary Statistics  
Sequencing vector: M13; M7815; 31% of reads  
Sequencing vector: Plasmid; n/a; 69% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 210748 bases at least Q40  
Consensus quality: 213655 bases at least Q30  
Consensus quality: 215058 bases at least Q20  
Insert size: 194000; agarose-fp  
Insert size: 216246; sum-of-contigs  
Quality coverage: 11.9 in Q20 bases; agarose-fp

Quality coverage: 10.7 in Q20.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 8149: contig of 8149 bp in length  
\* 8150 8249: gap of 100 bp  
\* 8250 9592: contig of 1343 bp in length  
\* 9593 9692: gap of 100 bp  
\* 9693 10733: contig of 1041 bp in length  
\* 10734 10833: gap of 100 bp  
\* 10834 13519: contig of 2686 bp in length  
\* 13520 13619: gap of 100 bp  
\* 13620 17510: contig of 3891 bp in length  
\* 17511 17610: gap of 100 bp  
\* 17611 24602: contig of 6992 bp in length  
\* 24603 24702: gap of 100 bp  
\* 24703 35434: contig of 10732 bp in length  
\* 35435 35534: gap of 100 bp  
\* 35535 124474: contig of 88940 bp in length  
\* 124475 124574: gap of 100 bp  
\* 124575 134664: contig of 10090 bp in length  
\* 134665 134764: gap of 100 bp  
\* 134765 162343: contig of 27579 bp in length  
\* 162344 162443: gap of 100 bp  
\* 162444 208917: contig of 46474 bp in length  
\* 208918 209017: gap of 100 bp  
\* 209018 217346: contig of 8329 bp in length.

## FEATURES

Source	Location/Qualifiers
1..217346	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="19"
	/map="19"
	/clone="RP11-795B6"
	/clone_lib="RPC1-11 Human Male BAC"
1..8149	/note="assembly_fragment"
clone_end:SP6	
vector_side:left	
8250..9592	/note="assembly_fragment"
9693..10733	/note="assembly_fragment"
10834..13519	/note="assembly_fragment"
13620..17510	/note="assembly_fragment"
17611..24602	/note="assembly_fragment"
24703..35434	/note="assembly_fragment"
35535..124474	/note="assembly_fragment"
124575..134664	/note="assembly_fragment"
134765..162343	/note="assembly_fragment"
162444..208917	/note="assembly_fragment"
209018..217346	/note="assembly_fragment"
clone_end:7	
vector_side:right	
BASE COUNT 55109 a 54154 c 53923 g 53053 t 1107 others	
ORIGIN	
Query Match	65.8%; Score 1316.2; DB 2; Length 217346;







clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

#### ----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 196656 bases at least Q40  
 Consensus quality: 197883 bases at least Q30  
 Consensus quality: 198879 bases at least Q20  
 Insert size: 215000; agarose-fp  
 Insert size: 198692; sum-of-contigs  
 Quality coverage: 9.03x in Q20 bases; agarose-fp  
 Quality coverage: 9.73x in Q20 bases; sum-of-contigs  
 -----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

```

1 28306: contig of 28306 bp in length
* 28307 28406: gap of unknown length
* 28407 37856: contig of 9450 bp in length
* 37857 37956: gap of unknown length
* 37957 73522: contig of 35566 bp in length
* 73523 73622: gap of unknown length
* 73623 83567: contig of 9945 bp in length
* 83568 83667: gap of unknown length
* 83668 88817: contig of 5150 bp in length
* 88818 88917: gap of unknown length
* 88918 125611: contig of 36694 bp in length
* 125612 125711: gap of unknown length
* 125712 159879: contig of 34168 bp in length
* 159880 159979: gap of unknown length
* 159980 174698: contig of 14719 bp in length
* 174699 174798: gap of unknown length
* 174799 186382: contig of 11584 bp in length
* 186383 186482: gap of unknown length
* 186483 193344: contig of 6862 bp in length
* 193345 193445: gap of unknown length
* 193445 199363: contig of 5919 bp in length
* 199364 199464: gap of unknown length
* 199464 200792: contig of 1329 bp in length.

```

#### FEATURES

```

source
1..200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-355A20"
/clone_lib="CH251"
1..28306
/notes="assembly_fragment"
clone_end:SP6
vector_side:left
28407..37856
/notes="assembly_fragment"
37957..73522
/notes="assembly_fragment"
73623..83567
/notes="assembly_fragment"
83668..88817
/notes="assembly_fragment"
88918..125611
/notes="assembly_fragment"
125712..159879

```

```

misc_feature
159980..174698
/notes="assembly_fragment"
misc_feature
174799..186382
/notes="assembly_fragment"
misc_feature
186483..193344
/notes="assembly_fragment"
misc_feature
193445..199363
/notes="assembly_fragment"
misc_feature
199464..200792
/notes="assembly_fragment"
clone_end:r7
vector_side:right"
BASE COUNT 49975 a 49805 c 48502 g 51409 t 1101 others
ORIGIN
Query Match 61.4%; Score 1227.2; DB 2; Length 200792;
Best Local Similarity 95.5%; Pred. No. 3.4e-292;
Matches 1323; Conservative 0; Mismatches 18; Indels 45; Gaps 4;
QY 654 GGCCATGATCATCTTCTTGACTATGCTGATGTGACAAAGTACCTAAAGCCCATCAGACTCTA 713
Db 78497 GTCTATGATCATCTTCTTGACTATGCTGATGTGACAAAGTACCTAAAGCCCATCAGACTCTA 78438
QY 714 CCCTTTAAATATGCAGTTTGGGCCAGGCACCGTGGCTCATGCTGTAAATCCAGCAGCTTT 773
Db 78437 CCCTTTAAATATGCAGTTTGGGCCAGGCACCTGTGGCTCATGCTGTAAATCCAGCAGCTTT 78378
QY 774 GGGAGCAGAGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGCAGCAGCCTGGCCCAACAT 833
Db 78377 GGGAGCAGAGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGCAGCAGCCTGGCCCAACAT 78318
QY 834 GGTGAACCTCTGCTTTTACTAAAAAATAAAAAAAAAAAAAAAAAAAAAATCAGCGGGTGTCTG 893
Db 78317 GGTGAACCTCTGCTTTTACT-----AAAAAATAAAAAAAAAAAAAATCAGCGGGTGTCTG 78264
QY 894 GGGGCACACTGTAAATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTTGAACCCCT 953
Db 78263 GGGGCACACTGTAAATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTTGAACCCCT 78204
QY 954 GGAGCGGAGGTTGCAGTGGCGGAGATCACATCACCGCCCTCCAGCCTGGGCGACAGAG 1013
Db 78203 GGAGGTGGAGTTGCAGTGGCGGAGATCACATCACCGCCCTCCAGCCTGGGCGACAGAG 78144
QY 1014 CAGACTCTGCTCAATAATAATAATAAACAACAGACAGCAGTGTGTTGACCTTAGT 1073
Db 78143 CAAGACTCTGCTCAATAATAATAATAAACAACAGACAGCAGTGTGTTGACCTTAGT 78084
QY 1074 TATATCT--AAAAAATAATGCTGTCAACAATAATAGACAGAGTGAATAAAGGAAATA 1131
Db 78083 TATACCTAAAAAATAATAATAATAATAAACAACAGACAGTGAATAAAGGAAATA 78024
QY 1132 AATGGCCCAAGAACTCTAAGGTATATTTGACAAATATTTCAGAACCTTTAAAAAGAAAG 1191
Db 78023 AATGGCCCAAGAACTCTAAGGTATATTTGACAAATATTTCAGAACCTTTAAAAAGAAAG 77964
QY 1192 AATCAGAGGAGATAGAAAGACAGGAGGACAGGAGACAGAAACACCTGTGCCCCAAG 1251
Db 77963 AATCAGAGGAGATAGAAAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 77904
QY 1252 GAGACAAAACAAAGGCTCTTAAGACAGACAGAGAG----GAGAGAGAGAGAGAGTGTAGTGA 1307
Db 77903 GAGACAAAACAAAGGCTCTTAAGACAGACAGGAGGAGGAGGAGAGAGAGAGAGAGTGTAGTGA 77844
QY 1308 GAGACAGACAGAGAAAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1347
Db 77843 GAGACAGACAGAGAAAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 77784
QY 1348 -----CACAGAGACAGAGAGGCGGAGAGGATAGAAAGAGAGAGAGAGAGGAGTGA 1394
Db 77783 AGAGAGAGGGGTGGAGAGAGACAGAGAGGCGGAGAGGAGGATAGAAAGAGAGAGAGAGGAGTGA 77724
QY 1395 GAGAGACACGAGATATTGAGAGAGAGACTCAGAAAGATAGCGGAGGAGAGACACAGAGAGAG 1454

```

Db 77723 GAGAGACGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGAGAAACACACAGAGA 77664  
|||||  
QY 1455 TGGAGAGACTCTGAGAAAAACCCAGACAAAGATGGAAGAGGAGTATCGAGGGTGA 1514  
|||||  
Db 77663 TGGAGAGACTCTGAAAAAACCCAGACAAAGATGGAAGAGGAGTATCGAGGGTGA 77604  
|||||  
QY 1515 ACACACAGTGTGGAATGAGCAAAATGACAGAAAGCAAGCAATCCAGGCGCCCAAGA 1574  
|||||  
Db 77603 ACACACAGTGTGGAATGAGCAAAATGACAGAAAGCAAGCAATCCAGGCGCCCAAGA 77544  
|||||  
QY 1575 ATAGTGACCCAGAGTGTGTGAGAGCCAGATCTTAAGGCTTGGGGGAGGAGGAGGAAGGG 1634  
|||||  
Db 77543 ATAGTGACCCAGAGTGTGTGAGAGCCAGATCTTAAGGCTTGGGGGAGGAGGAGGAAGGG 77484  
|||||  
QY 1635 CTGGCTTGGCTTCGGGAGACCCCTCCCATTTCTCCGGCCAGGAGGAGGTAGGAGTGACAT 1694  
|||||  
Db 77483 CTGGCTTGGCTTCGGGAGACCCCTCCCATTTCTCCGGCCAGGAGGAGGTAGGAGTGACAT 77424  
|||||  
QY 1695 TCCGACTGGGTGGGGGTGCTCTGGGGTGGAGATAGGGGAGCAGGAGGAGCTATTGC 1754  
|||||  
Db 77423 TCCGACTGGGTGGGGGTGCTCTGGGGTGGAGATAGGGGAGCAGGAGGAGCTATTGC 77364  
|||||  
QY 1755 TAAGCCCGATAGGACCTATTGCGCGGGAATGTGCCCGAGGAGCAGTGGGTGGTAT 1814  
|||||  
Db 77363 TAAGCCCGATAGGACCTATTGCGCGGGAATGTGCCCGAGGAGCAGTGGGTGGTAT 77304  
|||||  
QY 1815 AACTCAGGCGCGGTGCCAGAGCCAGGAGGAGCAGTGGCCAGGAGGACAGGCGCTGA 1874  
|||||  
Db 77303 AACTCAGGCGCGGTGCCAGAGCCAGGAGGAGCAGTGGCCAGGAGGACAGGCGCTGA 77244  
|||||  
QY 1875 GAAGTCTCGGCTGAGCTGGAGCAAAATCCCCACCCCTACCTGGGGGACAGGCGCAAGT 1934  
|||||  
Db 77243 GAAGTCTCGGCTGAGCTGGAGCAAAATCCCCACCCCTACCTGGGGGACAGGCGCAAGT 77184  
|||||  
QY 1935 GAGACCTGGTGGGTGGCTCAGCAGGAGGAGGAGGAGTGTCTGTGCTCTGTGACCC 1994  
|||||  
Db 77183 GAGACCTGGTGGGTGGCTCAGCAGGAGGAGGAGGAGTGTCTGTGCTCTGTGACCC 77124  
|||||  
QY 1995 CACATC 2000  
|||||  
Db 77123 CACATC 77118  
|||||  
RESULT 6  
AC130188/c  
LOCUS AC130188 176647 bp DNA linear HTG 14-NOV-2002  
DEFINITION Papio anubis clone RP41-421P3, WORKING DRAFT SEQUENCE, 12 ordered  
pieces.  
ACCESSION AC130188  
VERSION AC130188.2 GI:24960890  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
SOURCE Papio anubis (olive baboon)  
ORGANISM Papio anubis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
Cercopitheciinae; Papio.  
REFERENCE 1 (bases 1 to 176647)  
AUTHORS Akter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S. M.,  
Benjamin, B., Blakesley, R. W., Bouffard, G. G., Brinkley, C., Brooks, S.,  
Carliaga, K., Coleman, B., Engle, J., Granlitz, S., Guan, X., Gupta, J.,  
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J. R., Karlins, E.,  
Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, O. L., Maduro, V. B.,  
Marquilies, E. H., Mastello, C., Maskeri, B., McDowell, J.,  
Paquirigan, C., Pearson, R., Portnoy, M. E., Prasad, A.,  
Reddix-Dugue, N., Schandler, K., Schueler, M. G., Sison, C.,  
Stantripop, S., Thomas, J. W., Thomas, P. J., Touchman, J. W., Vogt, J. L.,  
Wetherby, K. D., Wiggins, L., Young, A. and Green, E. D.  
NISC Comparative Sequencing Initiative  
Unpublished  
REFERENCE 2 (bases 1 to 176647)  
AUTHORS Green, E. D.  
TITLE Direct Submission

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Submitted (08-AUG-2002) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
3 (bases 1 to 176647)  
Green, E. D.  
Direct Submission  
Submitted (14-NOV-2002) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
On Nov 14, 2002 this sequence version replaced gi:22138439.  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: [nisc\\_zoo@nhgri.nih.gov](mailto:nisc_zoo@nhgri.nih.gov)  
----- Project Information  
Center project name: dsr  
Center clone name: 421P03

The sequence data in this record represents an 'enhanced'  
version of a Phase 2 submission. Specifically, the indicated  
order and orientation of each sequence contig has been  
established using one or more of the following: read-pair  
data from individual subclones, overlaps with neighboring  
clones, alignment with available reference sequence (e.g.,  
human), and/or confirmation by PCR testing. In addition,  
the sequence assembly is based on at least 8x average  
coverage in Q20 bases and has been reviewed to rule out  
gross misassemblies, the low-quality ends of sequence  
contigs have been trimmed away, and each base is associated  
with a Phrap-derived quality score.

----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 171127 bases at least Q40  
Consensus quality: 173424 bases at least Q30  
Consensus quality: 174803 bases at least Q20  
Insert size: 188000; agarose-fp  
Quality coverage: 8.76x in Q20 bases; agarose-fp  
Quality coverage: 9.39x in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

\* 1  
\* 6527  
\* 6526: contig of 6526 bp in length  
\* 18812: contig of 12186 bp in length  
\* 18813  
\* 18912: gap of unknown length  
\* 18913  
\* 61780: contig of 42868 bp in length  
\* 61781  
\* 97068: contig of 35188 bp in length  
\* 97069  
\* 97168: gap of unknown length  
\* 97169  
\* 125206: contig of 28038 bp in length  
\* 125207  
\* 123306: gap of unknown length  
\* 123307  
\* 127021: contig of 1715 bp in length  
\* 127022  
\* 127121: gap of unknown length  
\* 127122  
\* 129387: contig of 2266 bp in length  
\* 129388  
\* 129487: gap of unknown length  
\* 129488  
\* 132302: contig of 2815 bp in length  
\* 132303  
\* 132402: gap of unknown length  
\* 132403  
\* 133735: contig of 3333 bp in length  
\* 135736  
\* 135836  
\* 143558: contig of 7723 bp in length  
\* 143559  
\* 143658: gap of unknown length  
\* 170182: contig of 26524 bp in length  
\* 170183  
\* 170282: gap of unknown length  
\* 176647: contig of 6365 bp in length.



```

SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Rump,A., Rosenthal,A., Drescher,B. and Schattevoy,R.
TITLE        Direct Submission
JOURNAL      Submitted (07-MAY-1997) Genome Analysis, Institute Of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE
AUTHORS      Lagemann,D. and Platzzer,M.
TITLE        Direct Submission
JOURNAL      Submitted (26-NOV-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE
AUTHORS      Lagemann,D. and Platzzer,M.
TITLE        Direct Submission
JOURNAL      Submitted (05-DEC-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
COMMENT
-----
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
-----
Project Information
-----
Center project name: B6
Center clone name: XX-540F24
-----
Summary Statistics
-----
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 32765 bases at least Q40
Consensus quality: 32770 bases at least Q30
Consensus quality: 32770 bases at least Q20
Quality coverage:13.12x
-----
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one pUC18 subclone.
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
FEATURES
source
-----
Location/Qualifiers
1..32799
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="xp22-PAR"
/clone="XX-B6cos"
537..642
/note="single stranded/single chemistry region"
misc_feature
2807..2820
/note="single stranded/single chemistry region"
misc_feature
2913..2927
/note="low quality region"
misc_feature
3166..3173
/note="single stranded/single chemistry region"
misc_feature
5131..5234
/note="single stranded/single chemistry region"
misc_feature
5235..5317
/note="single clone coverage"
-----
misc_feature
5266..5317
/note="low quality region"
misc_feature
5366..5431
/note="single clone coverage"
misc_feature
5432..5518
/note="single stranded/single chemistry region"
misc_feature
6151..6395
/note="single stranded/single chemistry region"
misc_feature
6303
unsure
6382
misc_feature
7106..7137
/note="single stranded/single chemistry region"
misc_feature
9265..9528
/note="single stranded/single chemistry region"
variation
9669
/note="T substituted in clone: XX-p99cos"
variation
10317
/replace="C"
/note="deleted in clone: XX-540F24"
/replace=""
10897
/note="G substituted in clone: XX-lBcos"
/replace="T"
11087..11090
/note="deleted in clone: XX-540F24"
/replace=""
11259
/note="T substituted in clone: XX-540F24"
/replace="C"
11744..11756
/note="low quality region"
11766..11880
/note="single stranded/single chemistry region"
11834
/note="T substituted in clone: XX-540F24"
/replace="C"
11892..11900
/note="deleted in clone: XX-lBcos"
/replace=""
12197..12200
/note="deleted in clone: XX-lBcos"
/replace=""
12256..12259
/note="deleted in clone: XX-lBcos"
/replace=""
12277
/note="A substituted in clone: XX-540F24"
/replace="C"
12290..12291
/note="deleted in clone: XX-540F24"
/replace=""
12312
/note="G substituted in clone: XX-540F24"
/replace="A"
13362
/note="G substituted in clone: XX-lBcos"
/replace="A"
13768
/note="T substituted in clone: XX-540F24"
/replace="G"
14138
/note="T substituted in clone: XX-540F24"
/replace="A"
14664
/note="C substituted in clone: XX-540F24"
/replace="T"
15355
/note="deleted in clone: XX-lBcos"
/replace=""
15447..15448
/note="deleted in clone: XX-lBcos"
/replace=""

```

variation	15481..15483 /note="deleted in clone: XX-540F24" /replace=""	15888	QY	625	TGGGATTACAGCGGTGAGCCACCGCGCCGCGCATGATCATCTTCTTACTATCTGATG 684 
variation	/note="A substituted in clone: XX-540F24" /replace="G"	15961	Db	6112	TGGGATTACAGCGGTGAGCCACCGCGCGCATGATCATCTTCTTACTATCTGATG 6053 
variation	/note="C substituted in clone: XX-540F24" /replace="T"	16252	QY	685	TGACAAAGTACCTTAAAGCCATCAGACTCTACCCCTTTAAATATGAGTGTGGCCAGGAC 744 
variation	/note="C substituted in clone: XX-540F24" /replace="T"	16334	Db	6052	ATTAAATTAAAGTGTGCGCGTGTGTTAGA-----AACTGGCAGCTTGTGGCCAGGACA 5999 
variation	/note="A substituted in clone: XX-1Bcos" /replace="G"	16584	QY	745	GTGGCTCATCCCTGTAATTCACGACTTTGGGAGGAGAGGTGGGTGAATCACTTGGCC 804 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	16793	Db	5998	GTGGCTCATCCCTGTAATTCACGACTTTGGGAGGAGAGGTGGGTGAATCACTTGGCC 5941 
variation	/note="G substituted in clone: XX-540F24" /replace="A"	17129	QY	805	CAGGAGTTTGAGACCCAGCTGGCCACACATGCTGCTCTTCTTCTTCTTCTTCTTCTTCT 864 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	17489	Db	5940	CAGGAGTTTGAGACCCAGCTGGCCACACATGCTGCTCTTCTTCTTCTTCTTCTTCTTCT 5891 
variation	/note="C substituted in clone: XX-540F24" /replace="G"	18712..18717	QY	865	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 924 
misc_feature	/note="single stranded/single chemistry region"	21509..21686	Db	5890	--AATATACAAAAATTAGCTGGGAGTGGTGGCGCATGCTTGTATATCCAGCACTTGGGA 5833 
variation	/note="A substituted in clone: XX-540F24" /replace="G"	22559..22586	QY	925	GGCTGAGGCAGAGAGTCACTTGAACCCCTGGAGCGGAGGTTCAGTGGCCGAGATCAC 984 
variation	/note="G substituted in clone: XX-540F24" /replace="T"	22606	Db	5832	GGCTGAGGCAGAGAGTCACTTGAACCCCTGGAGCGGAGGTTCAGTGGCCGAGATCAC 5773 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	22606	QY	985	ATCACCGCCCTCCAGCTGGCGGAGAGCAAGACTCTGCTCAATAATAAATAAATAAATA 1044 
variation	/note="C substituted in clone: XX-540F24" /replace="G"	22606	Db	5772	GCCTCTGCACTCCAGCT--GGTACAGAGCAAGACTCT-----TAAAGAGAGGAGG 5720 
misc_feature	/note="single stranded/single chemistry region"	22559..22586	QY	1045	AACGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1104 
variation	/note="A substituted in clone: XX-540F24" /replace="G"	22606	Db	5719	AAGGAAGGAAGAGAGAGGAG 5660 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	22606	QY	1105	AGAGCAG 1164 
variation	/note="G substituted in clone: XX-540F24" /replace="T"	22606	Db	5659	AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 5601 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	22606	QY	1165	ATCATTCAGAACCTTTAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1224 
variation	/note="G substituted in clone: XX-540F24" /replace="T"	22606	Db	5600	GAAAG 5541 
misc_feature	/note="single stranded/single chemistry region"	22559..22586	QY	1225	GGGAGACAGAACACCTGTGG--CCCAGGAGAGACAAACAAAGGCTCTCTAGACAGACAG 1283 
variation	/note="A substituted in clone: XX-540F24" /replace="G"	22606	Db	5540	AGGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5481 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	22606	QY	1284	AGGAG 1343 
variation	/note="G substituted in clone: XX-540F24" /replace="T"	22606	Db	5480	GGAGGAAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5421 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	22606	QY	1344	GAGACAG 1402 
variation	/note="G substituted in clone: XX-540F24" /replace="T"	22606	Db	5420	TGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 5361 
variation	/note="A substituted in clone: XX-540F24" /replace="G"	22606	QY	1403	CGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGAGAGAGAGAGAGAGAGAGAGAG 1462 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	22606	Db	5360	AAGAAGAAATGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 5301 
variation	/note="G substituted in clone: XX-540F24" /replace="T"	22606	QY	1463	GACTCTGAGAAACACAG 1522 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	22606	Db	5300	GAAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 5241 
variation	/note="A substituted in clone: XX-540F24" /replace="G"	22606	QY	1523	TGCTGGAATTCAGCAAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1582 
variation	/note="G substituted in clone: XX-540F24" /replace="T"	22606	Db	5240	AGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 5181 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	22606	QY	1583	CCAGAGTTGTGAG 1634 
variation	/note="G substituted in clone: XX-540F24" /replace="T"	22606	Db	5180	GGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 5129 

RESULT 8  
AC109997  
LOCUS  
DEFINITION

AC109997 141923 bp DNA linear HTG 09-FEB-2002  
Homo sapiens chromosome 5 clone CTD-207601, WORKING DRAFT SEQUENCE,





## Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

## Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

## BgIII

## HindIII

## EcoRI

SeqDerMap	FrngPrnt	SeqDerMap	FrngPrnt	SeqDerMap	FrngPrnt
2292	2308	6857	6790	8696	8687
2067	2005	6382	6450	6	<800
4763	4915	512	<800	3403	3464
2585	2578	449	<800	334	<800
1981	2005	7940	7924	951	930
5496	5398	10618	10372	4919	4873
8296	8303	305	<800	3694	3709
10488	10186	1052	1054	5455	5487
5655	5680	2795	2867	555	<800
6535	6574	1527	1521	13221	13066
2695	2731	9418	9389	8640	8687
10234	10186	2970	3015	5654	5681
4871	4915	10593	10372	2842	2844
1822	1781	891	883	10800	10593
726	720	1004	1054	11874	11745
540	<800	1340	1314	8529	8687
11711	11427	817	883	286	<800
25420	25600	4767	4687	6660	6724
2222	2308	7571	7564	13999	13754

360	<800	4492	4508	9092	9112
2344	2308	1092	1054	2356	2380
1198	1180	3332	3396	2287	2380
3828	3831	1123	1054	2669	2713
3328	3300	1803	1824	924	930
3248	3300	3041	3122	4641	4650
1289	1252	16514	16912	1247	1206
399	<800	192	<800	1708	1662
4134	4173	4763	4687	6082	6133
900	895	587	<800	1904	1897
3852	3963	5086	5041	4433	4394
4463	4378	2278	2421	6851	6941
3959	3831	1336	1314	1799	1783
1990	2005	3223	3122	23885	24086
4272	4173	7004	7057	-----	-----
2327	2308	971	1054	-----	-----
6858	6866	2341	2421	-----	-----
3056	3063	1649	1646	-----	-----
6815	6866	2822	2867	-----	-----
809	810	6084	6048	-----	-----
4433	4378	163	<800	-----	-----
6135	6032	9648	9389	-----	-----
-----	-----	311	<800	-----	-----
-----	-----	3279	3396	-----	-----
-----	-----	3900	3950	-----	-----
-----	-----	2035	2029	-----	-----
-----	-----	3612	3630	-----	-----
-----	-----	3053	3229	-----	-----
-----	-----	3414	3396	-----	-----
-----	-----	865	883	-----	-----
-----	-----	637	<800	-----	-----
-----	-----	259	<800	-----	-----
-----	-----	97	<800	-----	-----
-----	-----	1242	1314	-----	-----
-----	-----	270	<800	-----	-----
-----	-----	23	<800	-----	-----
-----	-----	47	<800	-----	-----







```

QY 746 TGCTCATGCTGTAAATCCAGCACTTTGGAGGCGAGAGTGGTGAATCACTTTGAGGCC 805
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91862 TGCTCAGCGCTGTATCCAGCACATTTGGAGGCCAGCCAGTGCATCACTGAGGTC 91921
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 806 AGAGATTGAGACCGCCTGGCCAAACATGGTGAACCTCTCTTTACTTAAAAA 865
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91922 CGAGTTTCGAGACCGAGCTGGCCAAAGCTGTCAAACCCCATCTCTACT----- 91969
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 866 AAAAAAATAATCAGCGGGTGTCTGGGCGCACACCTGTATCCAGCTATGCTGGAG 925
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91970 AAAAAATACAAAATAGCAGGCTTAGTGTGTGCTGTAGTCCAGCTACTCGGGAG 92029
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 926 GCTGAGGCGAGGAGTCATTAACCCCTGGAGCGGAGGTTCAGTGGCGGAGATCACA 985
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92030 TCTGAGGCGATGAGCAATCACTTGAACCTGGGAGGTTCAGTGAGCCAGATACG 92089
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 986 TCACCGCCCTCAGCTGGCGGACAGCAAGACTCTGTCTAAATAAATAAACA 1045
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92090 CCAGTGTACTCCAGCCTGGTGGACAGCAAGACTCTGTCTCAAAAAATAAATAA 92149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1046 ACGAA 1050
    || ||
Db 92150 AATAA 92154

```

## RESULT 12

```

AC098873
LOCUS AC098873 Homo sapiens BAC clone RP11-793H20 from 4, complete sequence.
ACCESSION AC098873 AC036224
VERSION AC098873.3 GI:18855169
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 119147)
AUTHORS Walligorski,J., Haakenson,W. and Spalding,L.
TITLE The sequence of Homo sapiens BAC clone RP11-793H20
JOURNAL Unpublished (2001)
MEDLINE 3 (bases 1 to 119147)
PUBMED 99063792
TITLE Waterston,R.H.
AUTHORS Waterston,R.H.
JOURNAL Submitted (04-NOV-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 119147)
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
TITLE Submitted (21-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 119147)
AUTHORS Waterston,R.
JOURNAL Direct Submission
TITLE Submitted (09-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb.21, 2002 this sequence version replaced gi:18042516.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0793H20
Drafting Center: WIBR

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

## SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTORS: pBACE3.6

## NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-260K18, 2000 bp overlap; the clone sequenced to the right is RP11-473N2. Actual start of this clone is at base position 106515 of RP11-260K18; actual end is at base position 119147 of RP11-793H20.

Single plasmid region exists between 77265 and 77307. Polymorphisms exist between AC079118, AC024191 and AC098873.

The sequence of AC036224 has been incorporated into AC098873.

FEATURES	source
Location/Qualifiers	1. .119147
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="4"
	/map="4"
	/clone="RP11-793H20"
	/clone_lib="RPCI-11"
repeat_region	167..273
repeat_region	/rpt_family="MaLR"
	841..887
	/rpt_family="T-rich"
misc_feature	870..889
	/note="match to EST BG114275 (NID:gl2607872)"
repeat_region	2009..2357
	/rpt_family="L2"
repeat_region	2369..2934
	/rpt_family="MaLR"
repeat_region	2959..3655
	/rpt_family="L2"
repeat_region	4049..4289
	/rpt_family="MIR"
repeat_region	4448..4566
	/rpt_family="MIR"
repeat_region	4908..4933
	/rpt_family="(TGAA)n"
repeat_region	4936..5061
	/rpt_family="MER103"
repeat_region	5574..5879
	/rpt_family="Alu"
repeat_region	5949..6101



## VERSION

AC079169.32 GI:21263154

## KEYWORDS

HTG.

## SOURCE

Homo sapiens (human)

## ORGANISM

## REFERENCE

## AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 108893)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,

Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbara,J.,

Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,

Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buha,C.,

Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,

Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,

Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,

Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,

Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,

Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,

Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,

Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,

Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,

Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,

Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Haves,A.,

Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,

Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,

Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,

Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,

Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,

Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,

Louiseged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,

Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,

Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,

Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,

Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,

Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,

Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,

Peery,J., Perez,L., Peters,D., Pickens,R., Primus,E., Pu,L.L.,

Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,

Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N.,

Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,

Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,

Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,

Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,

Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,

Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Woodden,S.,

Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,

Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 108893)

Worley,K.C.

Direct Submission

Submitted (23-AUG-2000) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 108893)

Worley,K.C.

Direct Submission

Submitted (29-MAY-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 108893)

Worley,K.C.

Direct Submission

Submitted (30-MAY-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 108893)

Worley,K.C.

Direct Submission

Submitted (01-JAN-2003) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

6 (bases 1 to 108893)

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Worley,K.C.

Direct Submission

Submitted (22-JAN-2003) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On May 30, 2002 this sequence version replaced gi:21206031.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email

gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the

entire insert of this clone. Overlapping regions of clones are only

sequenced and submitted once, so the sequence for the remainder of

the insert may be found in the record for the adjacent clones.

Overlapping clones are noted at the beginning and end of the

Features listing.

## ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches

of a local database that includes entries from dbSTS, GDB, and

local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,

unpublished) for Human and Mouse sequences.

Genes and region of sequence similarity are identified by BLAST

(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the

EST and cDNA sequences. Genes demonstrate at least two exons

flanked by consensus splice sites that maintained sequence

continuity across the splice junctions. Sequences that are not

identical matches are annotated as similar.

## SEQUENCING READ COVERAGE:

Sequencing is completed to a minimum

standard of double strand coverage with a minimum of 2 clones and 2

reads with no ambiguities or 2 chemistries with a minimum of 2

clones and 3 reads with no ambiguities, if the sequence quality for

a region does not meet this standard, it will be indicated in the

annotation as low coverage.

## QUALITY OF INDIVIDUAL BASES:

This sequence meets stringent quality

standards - estimated error rate less than 1 per 10,000 bases.

Reports of lowest quality individual bases and measures of base

quality are listed below. Description of the metrics can be found

at URL:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT:

## FEATURES

## Source

## misc\_feature

1..108893

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="X"

/clone="RP11-47911"

1..2067

/note="overlaps bases 108750..110816 of clone AC002404"

/function="clone overlap"

complement(549..580)

/rpt\_family="LIM4"

complement(581..871)

/rpt\_family="AluX"

complement(872..1369)

/rpt\_family="LIM4"

1370..1394

/rpt\_family="(TTTTG)n"

complement(1398..1675)

/rpt\_family="AluX"

complement(1760..2236)

/rpt\_family="LIM4b"

2259..2373

/rpt\_family="AluJb"

complement(2375..2450)

/rpt\_family="LIM4b"

complement(2472..2666)

/rpt\_family="LIMC/D"

2683..2963

repeat\_region

Db	50324	GAATGGTCTCGAATCTCCTGACCTGCTGATGATCCGCCCGCTCGCGCTCCCAAGTGTCTGGGA	50383
QY	630	TTACAGCGGTGAGCCACCGCGCCGATCATCTTCTTGACTATGCTGATGCTGACA	689
Db	50384	TTACAGCGGTGAGCCACCGCGCCGCTGCTTAAGTCTTTTATATGTTCTTGTGGAG	50443
QY	690	AGTACATAAGCCATCAGACTCTACCCCTTTAAATATGCAAGTTGGCGGACGACCGTGGC	749
Db	50444	AATATAT-TATTCAACATTCACCTTGTAATAGGAAATATATTTAGCCAGGAACAGTGGC	50502
QY	750	TCATGCTGTAATTCAGCAGCTTTGGGAGGACAGAGTGGTGAATCACTTGGAGCCAGGA	809
Db	50503	TCGCGCTTGTAATCCAGACCTTTGGGGGACAGAGGACGAGATCACTCTGAGTCAAGT	50562
QY	810	GTTTGTAGACAGCTGGCCCAACATGCTGAACTCTGCTTTTACTAAAAAATAAATAA	869
Db	50563	GTTTGTAGACAGCTGGCCCAACATGCTGAAATCCCATATCTACT-----AAAA	50610
QY	870	AAAAAATAATCAGCCGGGTGCTGTTGGGGCACACCTGTAATCCAGCTATGCTGGAGGCTG	929
Db	50611	ATACAAAAATTAGCTGGCGGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	50670
QY	930	AGGCAGCAGAGTCACTTTGAACCTTGAGGCGGAGGTTGCAGTGGCGGAGATCACATCAC	989
Db	50671	AGGCACAAGAATCAGCTTTGAACCCAGGAGGAGGTTGCAGTGGAGCCAGGATCGCGCT	50730
QY	990	CGCCCTCCAGCCTGGCGGACAGCAAGACTCTGCTCAAAATAAATAAATAAACA	1044
Db	50731	TGCACCTCCAGCCTGGGTGACCGAGATTCGCTCTCAAAAAAATAAATAAATAA	50785
RESULT 14			
AC135468			
LOCUS	AC135468	100208 bp	DNA linear PRI 17-MAY-2003
DEFINITION	Homo sapiens chromosome 1 clone RP11-298p9, complete sequence.		
ACCESSION	AC135468	AL450462	
VERSION	AC135468.3	GI:30841055	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z., Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and Haugen,E.D.		
TITLE	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 100208)		
AUTHORS	Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-OCT-2002)		Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
REFERENCE	3 (bases 1 to 100208)		
AUTHORS	Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z., Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and Haugen,E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-NOV-2002)		Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
REFERENCE	4 (bases 1 to 100208)		
AUTHORS	Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z., Saenphimmachak,C., Buckley,D., Kibukawa,M., Raymond,C. and Haugen,E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-MAY-2003)		Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
COMMENT	On May 17, 2003 this sequence version replaced gi:25705321.		
	----- Genome Center		
	Center: University of Washington Genome Center		
	Center Code: UWGC		
	Web site: <a href="http://www.genome.washington.edu">http://www.genome.washington.edu</a>		
	Contact: <a href="mailto:uwgchq@u.washington.edu">uwgchq@u.washington.edu</a>		

## Drafting Center: SC

----- Project Information  
 Center project name: chr-1  
 Center clone name: sc0392  
 sc0392 was derived from the original BAC (RP11-298P9)  
 using yeast homologous capture technique.

## ----- Summary Statistics

Sequencing vector: plasmid; 53% of reads  
 Sequencing vector: plasmid; L08752; 47% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 100206 bases at least Q40  
 Consensus quality: 100207 bases at least Q30  
 Consensus quality: 100208 bases at least Q20  
 Insert size: 100208; sum-of-contigs  
 Quality coverage: 25.1x in Q20 bases; sum-of-contigs

## Overlapping Sequences:

5': RP11-297123 (UWGC:sc0384) AC098656, 3777-bp overlap  
 3': RP5-871E2 AL161797, 3773-bp overlap

## Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.  
 All manually edited bases have been reduced to quality zero.  
 Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
 Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:  
 all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

## Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below.  
 The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC.  
 Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

HindIII				EcoRI				BglII			
SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt	SeqDerMap	FngrPrnt
27861	27870	2775	2822	4460	4442						
10835	10615	15690	15691	2067	2066						
8	<800	5046	4994	4849	4849						
1191	1145	593	<800	2389	2395						
541	<800	121	<800	8936	8933						
10071	10047	4176	4218	3336	3347						
7336	7338	5633	5550	859	854						
818	843	4670	4557	24	<800						
3168	3194	4598	4557	29	<800						

FEATURES  
 source

Location/Qualifiers  
 1..100208

8217	8355	1084	1065	1752	1748
7535	7745	77	<800	1057	1047
477	<800	766	776	297	<800
5023	5020	619	<800	10235	10177
61	<800	1381	1379	1580	1550
4008	4017	459	<800	336	<800
1890	1924	1679	1655	2820	2839
5242	5141	2384	2388	3407	3423
846	843	1873	1835	5948	5883
3480	3460	1659	1655	835	854
823	843	4721	4758	603	<800
2118	2082	3389	3384	4452	4442
1996	1924	1331	1379	6616	6652
2474	2515	2045	2050	1944	1944
2746	2745	1252	1237	10220	10177
		3321	3384	7959	7933
		1441	1379	3562	3548
		3094	3079	3091	3117
		3442	3471	9174	9315
		169	<800	5473	5486
		1055	1065	445	<800
		3927	3866		
		1288	1313		
		4579	4557		
		5480	5550		
		41	<800		
		5950	6031		
		2655	2688		
		15	<800		
		213	<800		
		432	<800		
		688	<800		
		172	<800		
		73	<800		
		2709	2745		

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-298P9"
/clone_lib="RPC1 human BAC library 11"
BASE COUNT 31168 a 20265 c 19464 g 29311 t
ORIGIN

Query Match 22.5%; Score 450; DB 9; Length 100208;
Best Local Similarity 81.2%; Pred. No. 2.3e-100;
Matches 599; Conservative 0; Mismatches 100; Indels 39; Gaps 5;

QY 344 CTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTAGAGCGAGTCGCTGTCGCCCGCAGG 403
Db 62188 CTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTAGAGCGAGTCGCTGTCGCCCGCAGG 62247

QY 404 CTGAGTGCAGTGGCGGATCTCGCTCACTGCAAGCTCCGCCCTCCCGGGTTACGCCAT 463
Db 62248 CTGAGTGCAGTGGCGGATCTCGCTCACTGCAAGCTCCGCCCTCCCGGGTTACGCCAT 62307

QY 464 TCTCTGCCTCAGCTCCCAAGTAGCTGGGACTACAGCGCCGCCACTAGCGCCGGCTA 523
Db 62308 TCTCTGCCTCAGCTCCCAAGTAGCTGGGACTACAGCGCCGCCACTAGCGCCGGCTA 62367

QY 524 ATTTTGTATTTTGTAGTAGACGGGTTTACCGGTTTAGCGGGATGGCCTCGATC 583
Db 62368 A-TTTTGTATTTTGTAGTAGACGGGTTTACCGGTTTAGCGGGATGGTCTCGATC 62426

QY 584 TCCTGACCTCGTATCGCGCCCGCTCGGCCCTCCCAAGTGTGGGATTACAGCGGTGAGC 643
Db 62427 TCCTGACCTCGTATCGCGCCCGCTCGGCCCTCCCAAGTGTGGGATTACAGCGGTGAGC 62486

QY 644 CACCGCCCGCGCC-----ATGATCACTCTTCTGACTATGCTGATGTCACAACTACATA 698
Db 62487 CACCGCCCGCGCCCTTTTTCCTTCTTAATCTCTATCTCTCTCAATATCTCAGTGC 62546

QY 699 AGCCATCAGACTCTACCCCTTT-----AAATATCAGCTTTGGG 735
Db 62547 ACTCCTTAGTCTTGACCTTTTCTGCCAGAAATAACAAAAGTGATAGATTCAGATGGG 62606

QY 736 CCAGCACCCTGGCTATGCTCTGTAATTCAGACTTTGGGAGCGCAGAGTGGGTGATC 795
Db 62607 CCGGTGGCGTGGTCTACACCTGTAATCCAGCACTTTGGGAGCGCTAGCAGCGGATC 62666

QY 796 ACTTGAGCCAGGAGTTTGACACAGCTGCGCCCAACATGTTGAACTCTGCTTTACTAA 855
Db 62667 ACTTGAGTCTAGGAGTTCGAGACCACTGCTGCCAATGAGAAACCTGCTCTACT-- 62724

QY 856 AAAAAAAAAAAAAAAAAAATCAGCCGGTGTGCTGGGGCACACCTGTAATCCCAGC 915
Db 62725 -----AAAAACATACAAAAATTAGCCAGCGCTGGTGGCACAGGCTGTAATCCCAGC 62777

QY 916 TATGCTGGAGCTGAGGCACAGAGTCACTTGAACCTTGGAGCGGAGGTGCAAGTGGCC 975
Db 62778 TACTAGGAGGCTGAGGAAAAAGATCACTTTGAATCTTGGAGCGGAGTTTGGCATGAGC 62837

QY 976 CGAGATCACATCACCGC--CTTCCAGCCTGGGCGACAGAGCAAGACTCTGTCTCAAAATAA 1034
Db 62838 CGAGATCACACTCTGCNACTCCAGCCTGGGCGACAGAGTCTGTCTCAAAAAA 62897

QY 1035 TAAATAAACAAACGAACA 1052
Db 62898 AAAAAAAAAAAAAAAAAAGA 62915

RESULT 15
AC092500
LOCUS AC092500 182569 bp DNA linear PRI 31-MAY-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-111P21, complete sequence.
ACCESSION AC092500 AC063916
VERSION AC092500.2 GI:21281542
KEYWORDS HTG.
```

```
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 182569)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 182569)
AUTHORS Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J., Ivey,R.G. and
Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (12-JUL-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
REFERENCE 3 (bases 1 to 182569)
AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.
and Haugen,E.D.
TITLE Direct Submission
JOURNAL Submitted (31-MAY-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On May 31, 2002 this sequence version replaced gi:14702054.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgchgsu.washington.edu
Drafting Center: BCM
----- Project Information
Center project name: chr-3
Center clone name: RP11-111P21 (bc0233)
----- Summary Statistics
Sequencing vector: M13; L08821; 46% of reads
Sequencing vector: plasmid; L08752; 54% of reads
Chemistry: Dye-primer Bodipy; 5% of reads
Chemistry: Dye-terminator ET; 36% of reads
Chemistry: Dye-terminator Big Dye; 59% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 182265 bases at least Q40
Consensus quality: 182511 bases at least Q30
Consensus quality: 182559 bases at least Q20
Insert size: 182569; sum-of-contigs
Quality coverage: 7.6x in Q20 bases; sum-of-contigs
-----
Overlapping Sequences:
5: RP11-107013 (UWGC:bc0229) AC099539, 25534-bp overlap
3: RP5-1053D16 (UWGC:bc0709) AC006515
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an
alternate chemistry or covered by high quality data (i.e., Phred
quality >= 30); an attempt was made to resolve all sequencing
problems, such as compressions and repeats; all regions were
covered by at least one plasmid subclone or more than one M13
subclone; and the assembly was confirmed by restriction digest.
-----
Sequence Validation:
This sequence has been validated by Multiple Complete Digest
fingerprinting. Comparison of the experimentally derived digest
fragments with sequence-predicted fragments is given below.
The electronically-digested sequence consists of both insert and
```



vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

HindIII			BglII			EcoRI		
SeqDerMap	FngrPrnt		SeqDerMap	FngrPrnt		SeqDerMap	FngrPrnt	
1769	1754		8212	8420		8696	8539	
6382	6517		2067	2185		6	<800	
512	<800		4790	4821		1332	1321	
449	<800		1012	1012		375	<800	
2329	2351		8184	8420		891	918	
864	896		446	<800		1804	1805	
86	<800		837	846		3923	3988	
9282	8651		6762	6845		2990	3041	
1607	1604		7177	7149		6498	6630	
3752	3762		612	<800		1141	1131	
5118	5056		3808	3700		6574	6630	
281	<800		361	<800		3033	3041	
7991	8049		6436	6411		140	<800	
2048	2091		1475	1444		59	<800	
5856	6034		9769	9480		140	<800	
3137	3149		4733	4653		5413	5707	
10058	10022		2135	2185		925	918	
10112	10022		930	920		11072	10825	
682	<800		1292	1312		1799	1805	
6527	6517		11292	11997		4025	3988	
212	<800		6438	6411		7257	7834	
855	896		14482	14593		1858	1805	
2126	2091		656	<800		10884	10825	
4504	4472		490	<800		1340	1321	
1479	1460		2794	2813		8413	8539	
266	<800		125	<800		18710	19330	
7408	7463		8400	8420		760	736	
4055	4006		3534	3453		1951	1936	
17535	17700		4480	4394		5919	5991	
52	<800		2214	2185		1538	1501	
8108	8651		8499	8420		7	<800	

	-----	954	896	2367	2378	703	<800
	-----	2319	2351	1845	1824	16180	15729
	-----	13494	13082	1386	1444	6022	5991
	-----	6106	6034	6231	6411	7893	7834
	-----	668	<800	2770	2813	2111	2134
	-----	2357	2351	1437	1444	311	<800
	-----	908	896	1306	1312	4499	4487
	-----	7455	7463	6874	6845	52	<800
	-----	9934	10022	2198	2185	3595	3595
	-----	8697	9220	1029	1012	335	<800
	-----	1118	1110	5322	5319	11268	10825
	-----	11816	12009	1336	1312	3058	3041
	-----			5424	5319	5716	5707
	-----		1339	1339	1312	5087	5418
	-----		15962	15847	4965		5175
FEATURES	Location/Qualifiers						
source	1. .182569						
	/organism="Homo sapiens"						
	/mol_type="genomic DNA"						
	/db_xref="taxon:9606"						
	/chromosomes="3"						
	/clone="RP11-111P21"						
	/clone_lib="RPCI human BAC library 11"						
BASE COUNT	54730 a 39871 c 37778 g 50190 t						
ORIGIN							
	Query Match 22.5%; Score 450; DB 9; Length 182569;						
	Best Local Similarity 80.6%; Pred. No. 2.5e-100;						
	Matches 566; Conservative 0; Mismatches 125; Indels 11; Gaps 3;						
Qy	351 TTTTNTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCCTGCTGTGCGCCAGCGTGGAGT 410						
Dd	36794 TTGTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCCTGCTGTGCGCCAGCGTGGAGT 36853						
Qy	411 GCAGTGGCGGAFTCCGGCTCACTGCAAGCTCCGCTCCCGGGTTCAGCCATTCTCCTG 470						
Dd	36854 GCAGTGGCGGAFTCCGGCTCACTGCAAGCTCCGCTCCCGGGTTCAGCCATTCTCCTG 36913						
Qy	471 CCTCAGCCTCCCAAAGTAGCTGGGACTACAGGCGCCGCCACTACGCCGGCTTAATTTTT 530						
Dd	36914 CCTCAGCCTACCAGTAGCTGGGACTACAGGCGCCGCCACTACGCCGGCTTAATTTTT 36973						
Qy	531 TGTAATTTTAGTAGACACGGGGTTTACCGTTTTAGCCGGGATGGCTCGATCTCCTGAC 590						
Dd	36974 GTATTTTGTAGTAGATGGGGTTTCAACGGTTTTAGCCGGGATGGTCTCGATCTCCTGAC 37033						
Qy	591 CTGCTGATCCGCCCGCTCGGCTCCCCAAGTCTGGGATTACAGGCTGAGCCACCGCG 650						
Dd	37034 CTGCTGATCCGCCCGCTCGGCTCCCCAAGTCTGGGATTACAGGCTGAGCCACCGCG 37093						
Qy	651 CCCGGCATGATCATCTCTTGACTATGCTGATGTGACAAGTACCTAAAGCCATCAGACT 710						
Dd	37094 CTTGGCCGAACAGCAATTTTTTTTTTCCCTAGGTTGATGTTCTGAAAATCACTTTTA 37153						
Qy	711 CTACCCCTTTAAATATGCAFTTTGGCCAGGCAACCGTGGCTCATGCTGTAAATCCAGCAC 770						
Dd	37154 CCTAAAGTGAATAA-----AGGGCTGGCGCAGTGGCTCATGCTATAATCCAGCAC 37206						



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2003, 20:39:54 ; Search time 493.736 Seconds  
(without alignments)  
10934.739 Million cell updates/sec

Title: US-09-936-271B-13\_COPY\_1\_2000

Perfect score: 2000

Sequence: 1 gggccagagtgaggcaag.....gtgctctgcaccacatc 2000

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2000	100.0	11570	21	Human KLK-L2 gene.
C 2	424.4	21.2	84607	20	Human PACAP genom
C 3	416.6	20.8	139904	24	Human cDNA differe
C 4	415.2	20.8	17397	22	Human cardiovascul
5	415.2	20.8	19334	22	Human cardiovascul
6	415.2	20.8	19345	22	Human cardiovascul
C 7	414.8	20.7	12970	22	Human reproductive
C 8	414.8	20.7	12970	23	Human testicular a

C 9	414.6	20.7	74037	24	ABK944412	DNA encoding endot
10	410.4	20.5	10445	22	AA103692	Human reproductive
11	406.2	20.3	12970	22	AA105001	Human reproductive
12	406.2	20.3	12970	23	ABL97894	Human testicular a
C 13	406	20.3	32224	22	AAK89986	Human digestive sy
14	405.8	20.3	4433	22	ABA18400	Human nervous syst
C 15	405.8	20.3	4433	22	AA540516	DNA encoding human
C 16	405.8	20.3	4433	22	AA104142	Human reproductive
17	404.6	20.2	33147	22	AAK67282	Human immune/haema
C 18	403.6	20.2	1503841	24	ABT00010	Human neuregulin 1
C 19	403.6	20.2	1503841	24	ABT01503	Human neuregulin 1
C 20	403.6	20.2	1503900	22	AAK95240	Human neuregulin-1
C 21	403.6	20.2	1503900	22	AAK96733	Human neuregulin-1
C 22	402.6	20.1	36785	22	AAK82208	Human immune/haema
C 23	401.8	20.1	4431	22	ABA18399	Human nervous syst
C 24	401.8	20.1	4431	22	AA540513	DNA encoding human
C 25	401.8	20.1	4431	22	AA104139	Human reproductive
C 26	401.6	20.1	23241	22	AAF97870	Human neuroblastom
C 27	401.6	20.1	23241	22	AAF97871	Human neuroblastom
C 28	400.6	20.0	17717	22	AAK82007	Human immune/haema
29	400.6	20.0	17717	22	AAK82008	Human immune/haema
30	400.6	20.0	17719	22	AAK82006	Human immune/haema
31	400	20.0	19038	22	AAF30492	Human PAK5 gene.
C 32	398.4	19.9	25971	22	AAK86336	Human immune/haema
C 33	396.6	19.8	37783	22	AAK70780	Human immune/haema
C 34	396.6	19.8	37783	22	AAK76625	Human immune/haema
C 35	396.6	19.8	37783	22	AAK80913	Human immune/haema
C 36	394.8	19.7	4033	22	AAF90327	Human JAFFA genom
C 37	394.4	19.7	3608	22	AAK67271	Human immune/haema
C 38	394.4	19.7	3608	22	AAK74891	Human immune/haema
C 39	394.4	19.7	3608	22	AAK83192	Human immune/haema
C 40	394.4	19.7	20247	22	AA136315	Human musculoskele
C 41	394.4	19.7	20247	25	ABX59303	CDNA encoding nove
C 42	393.4	19.7	109906	24	ABK94411	DNA encoding endot
C 43	393	19.7	139904	24	ABK83562	Human cDNA differe
C 44	392.8	19.6	222930	24	ABK84349	Human cDNA differe
C 45	391.4	19.6	19199	22	AAK70995	Human immune/haema

ALIGNMENTS

RESULT 1	
AAA95905	
ID	AAA95905 standard; DNA; 11570 BP.
XX	
AC	AAA95905;
XX	
DT	02-FEB-2001 (first entry)
XX	
DE	Human KLK-L2 gene.
XX	
KW	Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;
KW	Kallikrein-like protein; serine protease; cytosolic; cancer;
KW	prostrate cancer; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200053776-A2.
XX	
PD	14-SEP-2000.
XX	
PF	09-MAR-2000; 2000WO-CA00258.
XX	
PR	11-MAR-1999; 99US-0124260.
PR	01-APR-1999; 99US-0127386.
PR	21-JUL-1999; 99US-0144919.
XX	
PA	(MOUN ) MOUNT SINAI HOSPITAL.
XX	
PI	Yousef GM, Diamandis EP;
XX	
DR	WPI; 2000-587440/55.

DR P-PSDB; AAB21296.

PT New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L  
PT protein mediated disorders, especially cancer. -

XX  
protein mediated disorders, especially cancer.

PS Claim 1; Page 143-149; 184pp; English.

The present sequence is the coding sequence of the human KLK-L2 gene, which encodes a kallikrein-like protein. Kallikreins and kallikrein-like proteins are a subgroup of the serine protease enzyme family. They catalyse the selective cleavage of specific polypeptide precursors to release peptides with potent biological activity. Nucleic acids encoding kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used to treat conditions mediated by the kallikrein-like proteins.

Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 other;

Query Match	100.0%	Score 2000;	DB 21;	Length 11570;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2000;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	

Qy	1	GGCCCCAGAGTCAAGCCAGAGAGAGGAGTTCGAGAGCTCCCTCTGCAAAAGTGGCTTGAGTC	60
Db	1	GGCCCCAGAGTCAAGCCAGAGAGAGGAGTTCGAGAGCTCCCTCTGCAAAAGTGGCTTGAGTC	60
Qy	61	TCCCTCGCCTAAATTCAGGCGAGAGGAGGCAGAGAAACACAGGGAAGAGAGGGGTGGGG	120
Db	61	TCCCTCGCCTAAATTCAGGCGAGAGGAGGCAGAGAAACACAGGGAAGAGAGGGGTGGGG	120
Qy	121	AGAAAGAGAGAGAGAGAGAGAGAGAGAAATTAACACAACTACAGAAACACAGAGAGAACAC	180
Db	121	AGAAAGAGAGAGAGAGAGAGAGAGAGAGAAATTAACACAACTACAGAAACACAGAGAGAACAC	180
Qy	181	ACAGAGAGCCTGGGACACAGGACACACAGAGTTCAGAGAGAAAGAGAGAAGATAGAGAAAG	240
Db	181	ACAGAGAGCCTGGGACACAGGACACACAGAGTTCAGAGAGAAAGAGAGAAGATAGAGAAAG	240
Qy	241	ACACAAATGGAGACACAGAGGTGTAAGAAAGAGAGATTAAACAGAGTCCAGATACACGC	300
Db	241	ACACAAATGGAGACACAGAGGTGTAAGAAAGAGAGATTAAACAGAGTCCAGATACACGC	300
Qy	301	AAAGGGCAGAGACAGCTTTTCAGGCTGGTGCTATGATCATCTCTCTTTTTTTTTTTTT	360
Db	301	AAAGGGCAGAGACAGCTTTTCAGGCTGGTGCTATGATCATCTCTCTTTTTTTTTTTTT	360
Qy	361	TTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTCTGCCCGAGGTGGAGTGCAGTGGCGG	420
Db	361	TTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTCTGCCCGAGGTGGAGTGCAGTGGCGG	420
Qy	421	GATCTCGGCTCACTGCAAGCTCCGGCTCCCGGGTTACGCCATTCCTCGCTCAGCCTC	480
Db	421	GATCTCGGCTCACTGCAAGCTCCGGCTCCCGGGTTACGCCATTCCTCGCTCAGCCTC	480
Qy	481	CCAAGTAGCTGGGACTACAGGCGCGCGCACTACGCCCGGCTAATTTTTTGTATTTTA	540
Db	481	CCAAGTAGCTGGGACTACAGGCGCGCGCACTACGCCCGGCTAATTTTTTGTATTTTA	540
Qy	541	GTAGAGACGGGTTTACCGTTTTAGCCGGAGTGGCTCGATCTCTGACCTCGTGATCC	600
Db	541	GTAGAGACGGGTTTACCGTTTTAGCCGGAGTGGCTCGATCTCTGACCTCGTGATCC	600
Qy	601	GCCCGCTCGGCTCCCAAAGTGCTGGGATFACAGGCGTGAGCCACCGCGCCGGCCATG	660
Db	601	GCCCGCTCGGCTCCCAAAGTGCTGGGATFACAGGCGTGAGCCACCGCGCCGGCCATG	660
Qy	661	ATCATCTCTTGACTATGCTGATGTGCAAGTACCTTAAGGCCATFAGACTCTACCTTTA	720
Db	661	ATCATCTCTTGACTATGCTGATGTGCAAGTACCTTAAGGCCATFAGACTCTACCTTTA	720





KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
KW anti-infertility.  
XX Homo sapiens.  
XX WO200155321-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01340.  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 21-SEP-2000; 2000US-0234277.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236370.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.





PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225113.  
PR 14-AUG-2000; 2000US-0225114.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231142.  
PR 08-SEP-2000; 2000US-0231143.  
PR 08-SEP-2000; 2000US-0231144.  
PR 08-SEP-2000; 2000US-0231143.  
PR 08-SEP-2000; 2000US-0231141.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-0241826.  
PR 08-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.

PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-451930/48.

XX New cardiovascular system related polynucleotides and polypeptides,  
PT useful for diagnosing, treating and/or preventing disorders of the  
PT cardiovascular system -

XX Claim 1; SEQ ID NO 1943; 674pp; English.

XX Sequences AAS35741-AAS36942 represent genomic DNA molecules, which encode  
CC the cardiovascular system antigen polypeptides of the invention.  
CC Cardiovascular system antigens and their associated polynucleotides are  
CC useful in the diagnosis, treatment and prevention of various types of  
CC disorders in e.g. humans, mice, rabbits, goats, cats, dogs,  
CC chickens or sheep. A pathological condition can be determined by  
CC detecting the presence or absence of a mutation in a cardiovascular  
CC system antigen polynucleotide. The treatable disorders include autoimmune  
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such  
CC as neoplasms of the breast or liver, cardiovascular disorders such as  
CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,  
CC nervous system disorders such as Alzheimer's disease, infections caused  
CC by bacteria, viruses and fungi, ocular disorders such as corneal

CC infection, endocrine disorders such as premature labour and infertility,  
CC gastrointestinal disorders such as Crohn's disease, renal disorders such  
CC as glomerulonephritis and respiratory disorders such as asthma and  
CC pleurisy. The polypeptides can also be used to aid wound healing, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, to regenerate tissues and in chemotaxis.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

Query Match 20.8%; Score 415.2; DB 22; Length 19334;  
Best Local Similarity 78.6%; Pred. No. 5.1e-64;  
Matches 558; Conservative 0; Mismatches 123; Indels 29; Gaps 4;

```
QY 349 TTTTGTATTTTGTAGAGAGGGGTTTCTACGCTCAAGCTCCCGCTCCCGGGTTTCAGCCCATCTCTCC 468
Db 15143 TGTTCGTAACCTCTTTTGTGTGTAAAGACAGACTCTTGCTCTGCGCCAGGCTGGA 15202

QY 409 GTGCAGTGGGGGATCTCGGCTCACTGCAAGCTCCCGCTCCCGGGTTTCAGCCCATCTCTCC 468
Db 15203 GTGCAGTGGGACGATCTCGGCTCACTGCAAGCTCCCGCTCCCGGGTTTCAGCCCATCTCTCC 15262

QY 469 TGCCTCAGCTCCCAAGTAGCTGGGACTACAGGGCCCGGCTACCTACCGCCGGGTAAATTTT 528
Db 15263 TGCCTCAGCTCCCGAGTAGCTGGGACTACAGGCACCGCCGACCAAGCCCGGCTAA-TTT 15321

QY 529 TTTGTATTTTGTAGAGAGGGGTTTCTACGCTCAAGCTCCCGCTCCCGGGTTTCAGCCCATCTCTCC 588
Db 15322 TTTGTATTTTGTAGAGAGGGGTTTCTACGCTCAAGCTCCCGCTCCCGGGTTTCAGCCCATCTCTCC 15381

QY 589 ACCTCGTGATCCCGCCCGCTCCCGCTCCCGGGTTTCAGCCCATCTCTCC 648
Db 15382 ACCTTGATGATCGCCCGCTCCCGCTCCCGGGTTTCAGCCCATCTCTCC 15441

QY 649 CGCCCGGCGCATGATCTCTTCTGACTATGCTGATGTGAAGTACCAAGTACCAAGCCATCAGA 708
Db 15442 CACCCGGCC-----GATGTTTGAACCTCTTTAAGAGTGAACCAAAATCCCGAGCTTTTA-- 15494

QY 709 CTCCTACCTTTAAATATGATTTGGGCGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 768
Db 15495 -----TAAATATATAGCTGGGCTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 15545

QY 769 ACTTTGGAGGAGAGGTGGTGAATCACTTTGAGGCCAGGAGTTTGAGACCAGCTGCGCC 828
Db 15546 AGTTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 15605

QY 829 AACATGGTGAACCTCTCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 888
Db 15606 AACATGGTGAACCTCTCTTTTACT-----AAAAATACAAAAATTAAGCCAGGT 15653

QY 889 GTGCTGGGACACACCTCTGATCCAGCTATGCTGGAGGCTGAGGCGAGAGTCACTTGA 948
Db 15654 GTGCTGGGACACACCTCTGATCCAGCTATGCTGGAGGCTGAGGCGAGAGTCACTTGA 15713

QY 949 ACCCTGGAGGCGGAGGTGGAGTGGGCGGAGATCACATCCAGCCCTCCAGCCTGGGCGGA 1008
Db 15714 ACCCGGAGGCGGAGGTGGAGTGGGCGGAGATTCATCCAGCTATGCTGGAGGCTGAGGCGAGT 15773

QY 1009 CAGAGCAAGACTCTGCTTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1058
Db 15774 AAGAGCAAGACTCTGCTCAAAACATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15823
```

## RESULT 6

AAS36444

ID AAS36444 standard; DNA; 19345 BP.

XX AC AAS36444;

XX AC AAS36444;

DT 17-DEC-2001 (first entry)

XX Human cardiovascular system antigen genomic DNA SEQ ID No 1944.

XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
KW anti-infertility.  
XX  
OS Homo sapiens.  
XX  
PN WO200155321-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01340.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225477.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.



```
Db 15450 CACCCGGCC-----GATGTTTGAACCTCTTTAAACAGTGAACAAATCCACGCTTTTA-- 15502
QY 709 CTCCTACCTTTAAATATGCACTTTGGCCAGGACCGCTGCTCATGCTGTAAATCCAGC 768
Db 15503 -----TAAATATATAGCTGGCGCTGGCGCGGTGGCTCACACTGTAAATCCAGC 15553
QY 769 ACTTTGGAGGACAGAGTGGGTGAATCACCTTGAGGCCAGGAGTTTGAGACCAGCCTGGCC 828
Db 15554 AGTTTGGAGGACCAAGCAGGTGATCACCTGAGCTCGGGTGTTCAGACCAGCTTGGCC 15613
QY 829 AACATGTTGAACCTCTGCTTTACTATAAAAAAATAAAAAAATCAGCCGGT 888
Db 15614 AACATGTTGAACCTCTGCTTTACT-----AAAAATACAAAAATTAGCCAGT 15661
QY 889 GTCGTGGGACACCTCTGTAATCCAGCTATGCTGGAGGCTGAGGCAGAGTCACTTGA 948
Db 15662 GTGTGGCACACACCTGTAATCCAGCTACTTGGGAGGCTGAAGCAGAGTAATCACTTGA 15721
QY 949 ACCCTGGAGGCGAGGTGGCAGTGGCGCGAGATCACATCACCGCCCTCCAGCCTGGCGGA 1008
Db 15722 ACCCGGAGGCGAGGTGGCAGTGGCGCGAGATTATGCCACTGCACCTCCACTCTGGGTGA 15781
QY 1009 CAGAGCAAGACTCTGCTCAATAAATAAATAAACAACGAACAGCAGT 1058
Db 15782 AAGAGCAAGACTCTGCTCAAAACATTAATAAATAAATAAATAAAGT 15831

RESULT 7
AAL05001/c
ID -AAL05001 standard; DNA; 12970 BP.
XX
AC AAL05001;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 7689.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
XX cancer; gene therapy; ds.
XX Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
```

08-NOV-2000; 2000US-0246527.  
 08-NOV-2000; 2000US-0246528.  
 08-NOV-2000; 2000US-0246532.  
 08-NOV-2000; 2000US-0246609.  
 08-NOV-2000; 2000US-0246610.  
 08-NOV-2000; 2000US-0246611.  
 08-NOV-2000; 2000US-0246613.  
 08-NOV-2000; 2000US-0249207.  
 08-NOV-2000; 2000US-0249208.  
 08-NOV-2000; 2000US-0249209.  
 08-NOV-2000; 2000US-0249210.  
 08-NOV-2000; 2000US-0249211.  
 08-NOV-2000; 2000US-0249212.  
 08-NOV-2000; 2000US-0249213.  
 08-NOV-2000; 2000US-0249214.  
 08-NOV-2000; 2000US-0249215.  
 08-NOV-2000; 2000US-0249216.  
 08-NOV-2000; 2000US-0249217.  
 08-NOV-2000; 2000US-0249218.  
 08-NOV-2000; 2000US-0249244.  
 08-NOV-2000; 2000US-0249245.  
 08-NOV-2000; 2000US-0249264.  
 08-NOV-2000; 2000US-0249265.  
 08-NOV-2000; 2000US-0249297.  
 08-NOV-2000; 2000US-0249299.  
 08-NOV-2000; 2000US-0249300.  
 08-NOV-2000; 2000US-0250160.  
 08-NOV-2000; 2000US-0250391.  
 08-NOV-2000; 2000US-0251030.  
 08-NOV-2000; 2000US-0251988.  
 08-NOV-2000; 2000US-0256719.  
 08-NOV-2000; 2000US-0251479.  
 08-NOV-2000; 2000US-0251856.  
 08-NOV-2000; 2000US-0251868.  
 08-NOV-2000; 2000US-0251869.  
 08-NOV-2000; 2000US-0251989.  
 08-NOV-2000; 2000US-0251990.  
 08-NOV-2000; 2000US-0254097.  
 08-NOV-2000; 2000US-0259678.  
 (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;  
 WPI; 2001-465570/50.

Isolated nucleic acid molecule encoding a reproductive system antigen  
 is used in preventing, treating or ameliorating a medical condition -

Disclosure; SEQ ID NO 7689; 1297pp + Sequence Listing; English.

The present invention provides the protein and coding sequences of a  
 number of human reproductive system related antigens. These can be used  
 in the prevention and treatment of reproductive system disorders,  
 including cancer. The present sequence is a genomic sequence encoding a  
 protein of the invention.

Sequence 12970 BP; 3310 A; 2917 C; 2938 G; 3805 T; 0 other;

Query Match 20.7%; Score 414.8; DB 22; Length 12970;  
 Best Local Similarity 73.3%; Pred. No. 5.8e-64;  
 Matches 608; Conservative 0; Mismatches 197; Indels 25; Gaps 5;

358 TTTTCTTTTCTTTTCTTTGAGACGGAGTCTCTGTCGCCAGGCTGGAGTGGAGTGG 417  
 6179 TTTTCTTTTCTTTTCTTTGAGACGGAGTCTCTGTCGCCAGGCTGGAGTGGAGTGG 6120

418 CGGGATCTCGGCTACTGCAAGCTCCGGCTCCGGGTTTACGCCATTCTCTGCTCAGC 477  
 6119 CGTATCTCGGCTACTGCAAGCTCTCTGTTGGTTCACGCCATTCTCTGCTCAGC 6060

478 CTCCCAAGTAGCTGGGACTACAGGCGCCGCGCTACGCGGCTAAATTTTGTGATTT 537  
 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT

Db 6059 CTCCCGAGTAGCTGGGACTACAGGGCCCTGCCACCATGCCCGCTAAATTTTTT --TATTT 6002  
 QY 538 TTASTAGACAGCGGTTTTCACCGTTTACCGCGGATGGCTCGATCTCTGACCTGCTGA 597  
 Db 6001 TTAGTAGACACAGGTTTTCACCGTTTACCGCGGATGGTCTCCATCTCTGACCTCTTGA 5942  
 QY 598 TCCGCGCGCTCGGCTCCCAAAGTCTCGGATTTACAGGCGTGAGCCACCGCGCGCGCC 657  
 Db 5941 TCTGCTCGCTCGGCTCCCAAAGTCTCGGATTTACAGGCGTGAGCCACTGTGCGCGACC 5882  
 QY 658 ATGA-TCATCTCTTGTGACTATCTGATGTGACAAGTACCTAAAGCCATCAGACTTACCC 716  
 Db 5881 ACCACTCAGTGTTTTAAATGGTAAAAATATTGTACAGATACTTTAAACAAC----- 5829  
 QY 717 TTTAAATATGCAAGTTTGGGCCAGCAGCGTGCATGCTCATGCTCTTAATTCAGCACTTTGG 776  
 Db 5828 -GAAATATGTGGCGCGGCCAGCGCGTGCCTACGCCCTGTATATCCAGCACTTTGG 5770  
 QY 777 AGCAGAGTGGTGAATCAGTTGAGCGCAGAGTGTGAGACCCAGCTTGGCCAAACATGTT 836  
 Db 5769 AGGCCGAGTGGCAGATCAC--GAGGTCAAGAGATCGAGACCATCTGCTTAACACGTT 5712  
 QY 837 GAACTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 896  
 Db 5711 GAAACCTCTCTACTAAAAAT-----ATAAAAAATTAGCCATGCGTGTGTGAC 5664  
 QY 897 GCACACCTGTAATCCAGCTATGCTGGAGGCTGAGGCAGAGAGTACCTTTGAACCTTGA 956  
 Db 5663 AGGCACCTGTAGTCCAGCTACTTGGGAGGCTGTATGAGGAGAAATTTCCAAACCCGGGA 5604  
 QY 957 GCGGAGTGTGAGTGGCGGAGATCATATCACCGCTTCCAGCCTGGCGGACAGACAA 1016  
 Db 5603 GGTGGAGGCTGAGTGGCGGAGATCATATCACCGCTTCCAGCCTGGCGGACAGACAA 5544  
 QY 1017 GACTCTGCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1076  
 Db 5543 GACTCCATCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5484  
 QY 1077 ATCTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1136  
 Db 5483 ATGCACATGAAAGAAGCTTTAGTATCAGAAAAATGGAATTTAAACGACATGAATA 5424  
 QY 1137 GCCAAGACTTAAGGTATATTTGACAAATCATTTAGAACCTTTAAAAAA 1186  
 Db 5423 CCATACACACCCAGTAGAATGCTTAAATTAAGAACACTGACAAACAA 5374

## RESULT 8

ABL97894/c

ID ABL97894 standard; DNA; 12970 BP.

XX ABL97894;

AC ABL97894;

XX 21-JUN-2002 (first entry)

XX Human testicular antigen encoding DNA fragment SEQ ID NO: 2546.

XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;

KW reproductive system disorder; urinary system disorder; gene therapy;

KW cardiovascular disorder; respiratory disorder; neurological disorder;

KW gastrointestinal disease; infection; cytostatic; gene; ds.

OS Homo sapiens.

XX WO20015317-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01329.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483232/52.

Nucleic acids encoding 973 human testicular antigen polypeptides,  
useful for preventing, diagnosing and/or treating testicular cancer

Disclosure; SEQ ID NO 2546; 766pp; English.

The present invention provides the protein and coding sequences of 973

CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer.  
CC especially testicular cancers. The present sequence is a DNA encoding a  
CC protein fragment of the invention.

XX  
SQ Sequence 12970 BP; 3310 A; 2917 C; 2938 G; 3805 T; 0 other;

Query Match 20.7%; Score 414.8; DB 23; Length 12970;  
Best Local Similarity 73.3%; Pred No. 5.8e-64;  
Matches 608; Conservative 0; Mismatches 197; Indels 25; Gaps 5;

QY 358 TTTTCTTTTCTTTTGTGAGCGGAGTCTCTCTCTGTCGCCAGGCTGGAGTGCAGTGG 417  
DB 6179 TTTTCTTTTCTTTTGTGAGCGGAGTCTCTCTCTGTCGCCAGGCTGGAGTGCAGTGG 6120

QY 418 CGGGATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCACGCCAATTCCTCGCTCAGC 477  
DB 6119 CGTATCTCGGCTCACTGCAAGCTCCACTTCTGGGTTTCACGCCAATTCCTCGCTCAGC 6060

QY 478 CTCCTCAAGTAGCTGGGACTACAGGCGCGCCACTACGCCGCTAAATTTTCTTATTT 537  
DB 6059 CTCCTCAAGTAGCTGGGACTACAGGCGCGCTCCCTGCGGCTAAATTTTCTTATTT 6002

QY 538 TTAGTAGACGGGCTTTCACCGTTTGTAGCCGGGATGGCTCGATCTCTCGACCTCGTGA 597  
DB 6001 TTAGTAGACACAGGCTTTCACCGTTTGTAGCCAGGATGGTCTCTCTCGACCTCTGA 5942

QY 598 TCCGCCCGCTCGGCTCCCAAGTGTGGGATTACAGCGTGGAGCCAGCCGCGCCGCCGCC 657  
DB 5941 TCTGCTCGCTCGGCTCCCAAGTGTGGGATTACAGCGTGGAGCCAGCTGTGCCGCCGCC 5882

QY 658 ATGA-TCACTCTTCTGACTATCTGATGTGCAAGTACTTAAGCCATCAGACTCTACCC 716  
DB 5881 ACCACTCAGTGTTTTAATGGTAAATAATTTGTACAGATACCTTAACAAC----- 5829

QY 717 TTTAATATGCAAGTTTGGCCAGGCGGCTGCTCATGCTGTAATTCAGACCTTTGGG 776  
DB 5828 -GAAATATGTGGCGGCGGCGGCGGCTGCTCATGCTGTAATTCAGACCTTTGGG 5770

QY 777 AGGACAGAGTGGTGAATCACTTGAAGCGAGAGTTTGAGACCGCTGGCCCAACATGGT 836  
DB 5769 AGGCGGAGTGGGCGAGATCAC--GAGGTCAAGAGATCGAGACCATCTGGCTTAACACGGT 5712

QY 837 GAACTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 896  
DB 5711 GAAACCTGCTCTCTAAATAA-----ATAAATAAATAAATAAATAAATAAATAAATAA 5664

QY 897 GCACACCTGTAATCCAGCTATGCTGGAGGCTGAGGCGACGAGATCACTTGAACCTTGA 956  
DB 5663 AGGCACCTGTAGTCCAGCTACTTGGAGGCTGATGACGAGAAATGCTCAACCCGGGA 5604

QY 957 GCGGAGTGTGAGTGGCGGAGATCAATCAACCGCTTCCAGCTGGCGGACAGAGAA 1016  
DB 5603 GGTGGAGGCTGAGTGGCGGAGATCAATCAACCGCTTCCAGCTGGCGGACAGAGAA 5544

QY 1017 GACTCTGCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1076  
DB 5543 GACTCCATCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5484

QY 1077 ATCTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1136  
DB 5483 ATGCACATGAAGAAGCTTTAGTCAATCAGAAAAATGAAATTAAGACACATGAATA 5424

QY 1137 GCCAAGACTCTAAGTATATTGCAATCAATCAGAACCTTTAAAAA 1186  
DB 5423 CCAATACACACCCAGTAGAATGGCTAAATAAATAAATAAATAAATAAATAAATAA 5374

RESULT 9  
ABK94412/c  
ID ABK94412 standard; DNA; 74037 BP.

XX  
AC ABK94412;  
XX  
DT 27-AUG-2002 (first entry)  
XX  
DE DNA encoding endothelin converting enzyme 1 (ECE-1) #2.  
XX  
KW Endothelin; EDN; endothelin converting enzyme 1; ECE 1; EDNR;  
KW endothelin receptor; signaling system; cardiovascular disease;  
KW coronary heart disease; hypertension; atherosclerosis; angiogenesis;  
KW fatty acid metabolism; diabetes; familial hypercholesterolaemia;  
KW forensic marker; transgenic animal; solid support; SNP;  
KW cardiovascular regulator; gene; ds; single nucleotide polymorphism.  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT variation replace(8625,T)  
FT /\*tag= a  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(8938,T)  
FT /\*tag= b  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(9351,G)  
FT /\*tag= c  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(9439,C)  
FT /\*tag= d  
FT /\*standard\_name= "Single nucleotide polymorphism"  
XX  
XX WO200224747-A2.  
XX  
XX  
PD 28-MAR-2002.  
XX  
XX 31-AUG-2001; 2001WO-BP10087.  
XX  
XX 19-SEP-2000; 2000EP-0120123.  
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
XX  
XX Brinkmann U, Hoffmeyer S;  
XX  
XX WPI; 2002-435060/46.  
XX  
XX Novel polynucleotide of the endothelin/endothelin converting  
XX enzyme/receptors of endothelin and endothelin converting enzyme  
XX signaling system associated with cardiovascular disease, useful for  
XX treating the disease  
XX  
XX Claim 1; Page -: 190pp; English.  
XX  
XX The invention describes a polynucleotide (I) of the endothelin  
XX (EDN)/endothelin converting enzyme (ECE)/receptors of EDN and ECE (EDNR)  
XX signaling system which is associated with a cardiovascular disease. (I),  
XX the gene encoding EDN, ECE or EDNR (II) or a vector (III) expressing (I)  
XX or (II) is useful for producing cells capable of expressing a molecular  
XX variant polypeptide which is associated with a cardiovascular disease.  
XX (II), (III), the EDN, ECE or EDNR polypeptide, or a cell expressing  
XX a molecular variant gene comprising (I) is useful for identifying and  
XX obtaining a pro-drug or drug capable of modulating the activity of a  
XX molecular variant of a polypeptide of the EDN/EDNR/ECE signaling system  
XX or its gene product, or for identifying and obtaining an inhibitor of  
XX the activity of a molecular variant of a polypeptide of the EDN/EDNR/ECE  
XX signaling system or its gene product. The isolated proteins and  
XX polynucleotides encoding them are useful for preparation of a  
XX pharmaceutical composition for treating a cardiovascular disease such as  
XX coronary heart disease, hypertension, atherosclerosis, or related to  
XX abnormal angiogenesis or fatty acid metabolism e.g. diabetes and familial  
XX hypercholesterolaemia. The gene or a polynucleotide fragment of the  
XX EDN/ECE/EDNR signaling system are useful as forensic markers, for  
XX creating a transgenic animal and in creation of a solid support  
XX comprising polynucleotides, genes, vectors, polypeptides, antibodies or  
XX host cells of the invention. This sequence encodes a fragment of the





PR 27-SEP-2000: 2000US-0235836.  
PR 29-SEP-2000: 2000US-0236327.  
PR 29-SEP-2000: 2000US-0236367.  
PR 29-SEP-2000: 2000US-0236368.  
PR 29-SEP-2000: 2000US-0236369.  
PR 29-SEP-2000: 2000US-0236370.  
PR 02-OCT-2000: 2000US-0236370.  
PR 02-OCT-2000: 2000US-0236802.  
PR 02-OCT-2000: 2000US-0237037.  
PR 02-OCT-2000: 2000US-0237038.  
PR 02-OCT-2000: 2000US-0237039.  
PR 02-OCT-2000: 2000US-0237040.  
PR 13-OCT-2000: 2000US-0239935.  
PR 13-OCT-2000: 2000US-0239937.  
PR 20-OCT-2000: 2000US-0240960.  
PR 20-OCT-2000: 2000US-0241221.  
PR 20-OCT-2000: 2000US-0241785.  
PR 20-OCT-2000: 2000US-0241786.  
PR 20-OCT-2000: 2000US-0241787.  
PR 20-OCT-2000: 2000US-0241808.  
PR 20-OCT-2000: 2000US-0241809.  
PR 20-OCT-2000: 2000US-0241826.  
PR 01-NOV-2000: 2000US-0244617.  
PR 08-NOV-2000: 2000US-0246474.  
PR 08-NOV-2000: 2000US-0246475.  
PR 08-NOV-2000: 2000US-0246476.  
PR 08-NOV-2000: 2000US-0246477.  
PR 08-NOV-2000: 2000US-0246478.  
PR 08-NOV-2000: 2000US-0246523.  
PR 08-NOV-2000: 2000US-0246524.  
PR 08-NOV-2000: 2000US-0246525.  
PR 08-NOV-2000: 2000US-0246526.  
PR 08-NOV-2000: 2000US-0246527.  
PR 08-NOV-2000: 2000US-0246528.  
PR 08-NOV-2000: 2000US-0246529.  
PR 08-NOV-2000: 2000US-0246532.  
PR 08-NOV-2000: 2000US-0246610.  
PR 08-NOV-2000: 2000US-0246611.  
PR 08-NOV-2000: 2000US-0246613.  
PR 17-NOV-2000: 2000US-0249207.  
PR 17-NOV-2000: 2000US-0249208.  
PR 17-NOV-2000: 2000US-0249209.  
PR 17-NOV-2000: 2000US-0249210.  
PR 17-NOV-2000: 2000US-0249211.  
PR 17-NOV-2000: 2000US-0249212.  
PR 17-NOV-2000: 2000US-0249213.  
PR 17-NOV-2000: 2000US-0249214.  
PR 17-NOV-2000: 2000US-0249215.  
PR 17-NOV-2000: 2000US-0249216.  
PR 17-NOV-2000: 2000US-0249217.  
PR 17-NOV-2000: 2000US-0249218.  
PR 17-NOV-2000: 2000US-0249244.  
PR 17-NOV-2000: 2000US-0249245.  
PR 17-NOV-2000: 2000US-0249264.  
PR 17-NOV-2000: 2000US-0249265.  
PR 17-NOV-2000: 2000US-0249297.  
PR 17-NOV-2000: 2000US-0249299.  
PR 17-NOV-2000: 2000US-0249300.  
PR 01-DEC-2000: 2000US-0250160.  
PR 01-DEC-2000: 2000US-0250391.  
PR 05-DEC-2000: 2000US-0251030.  
PR 05-DEC-2000: 2000US-0251988.  
PR 05-DEC-2000: 2000US-0256719.  
PR 08-DEC-2000: 2000US-0251479.  
PR 08-DEC-2000: 2000US-0251856.  
PR 08-DEC-2000: 2000US-0251868.  
PR 08-DEC-2000: 2000US-0251869.  
PR 08-DEC-2000: 2000US-0251989.  
PR 08-DEC-2000: 2000US-0251990.  
PR 11-DEC-2000: 2000US-0254097.  
PR 05-JAN-2001: 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465570/50.  
DR Isolated nucleic acid molecule encoding a reproductive system antigen  
XX PT is used in preventing, treating or ameliorating a medical condition -  
PT Disclosure; SEQ ID NO 6380; 1297pp + Sequence Listing; English.  
XX PS  
XX CC The present invention provides the protein and coding sequences of a  
XX CC number of human reproductive system related antigens. These can be used  
XX CC in the prevention and treatment of reproductive system disorders,  
XX CC including cancer. The present sequence is a genomic sequence encoding a  
XX CC protein of the invention.  
SQ Sequence 10445 BP; 2469 A; 2885 C; 2678 G; 2413 T; 0 other;  
  
Query Match 20.5%; Score 410.4; DB 22; Length 10445;  
Best Local Similarity 78.9%; Pred. No. 3.4e-63;  
Matches 560; Conservative 0; Mismatches 126; Indels 24; Gaps 5;  
  
QY 359 TTTT TTTT TTTT TTTT TGACGAGTCTCGCTGTCGCCAGGCTGGAGTGCAGTGGC 418  
DB 8627 TTTT TTTT TTTT TTTT TGACGAGTCTCGCTGTCGCCAGGCTGGAGTGCAGTGGC 8686  
  
QY 419 GGGATCTGGCTCACTGCAAGCTCCGCTCCCGGGTTTCAGGCCATTCTCTGCCTCAGCC 478  
DB 8687 GCTATCTGGCTCACTGCAAGCTCTGCTCCCGGGTTTCAGGCCATTCTCTGCCTCAGTC 8746  
  
QY 479 TCCCAAGTAGCTGGGACTACAGGCGCCGCTACGCTACGCGCGGCTAAATTTTGTATTTT 538  
DB 8747 TCCCAAGTAGCTGGGACTACAGGCGCCGCTACGCTACGCGCGGCTAAATTTTGTATTTT 8805  
  
QY 539 TAGTAGAGACGGGTTTCACCGTTTAGCCGGGATGGCTCGATCTCTGACCTCGTGTAT 598  
DB 8806 TAGTAGAGACGGGTTTCACCGTTTAGCCGGGATGGCTCGATCTCTGACCTCGTGTAT 8865  
  
QY 599 CCGCCGCTCGGCTCCCAAGTGTGGGATTACAGCGGTACAGCCGCTAGCCGCGCGGCCA 658  
DB 8866 CCACCGCTCAGCTCCCAAGTGTGGGATTACAGCGGTACAGCCGCTAGCCGCGCGGCC 8925  
  
QY 659 TGATCATCTTCTGACTATGCTGATGTGACAAGTACCTAAGCCATCAGACTCTACCCCTT 718  
DB 8926 AGAATCTGCATTTTAAAAAATCCCAAGTCTTAAGTACAGGGGCATCTTAG----- 8976  
  
QY 719 TAAATATGCACTTTGGCCAGGACCGTGTCTCATGCTGTAATTCAGCAGCTTTGGAG 778  
DB 8977 AATAAGAAATTTAGGCTGGGCTGGTGTCTCAAGCTCTAATCTCAGCA-TTTGGGAG 9035  
  
QY 779 GCAGAGTGGTGAATCACTTGAGGCCAGGAGTTTGAGACAGCTTGCCCAACATGTTGA 838  
DB 9036 GCGGGGTGGCAGATTACCTGGGTGAGGTTCAAGACAGCAGCTGGTCAACATGTTGA 9095  
  
QY 839 AACTCTGTCTTTACTAAAAAATAAAAAAATAAAAAAATAACGCGGGTGTCTGGGGC 898  
DB 9096 AACCCCATCTGTACT-----AAAAACAATAAAATTTGCTGCCATCTGTTGG 9144  
  
QY 899 ACACCTGTATCCAGCTATGCTGGAGGCTGAGGCACGAGATCACTTGAACCTTGGAGG 958  
DB 9145 GTACCTTTAATCCAGCTACTCAGGAGGCTGAGCAGGAAATCGTTGAACCTTGGAGG 9204  
  
QY 959 CGGAGGTTGCAGTGGGCGGAGATCACAATCAGCCGCTCCAGCTTGGCGGAC--AGAGCAA 1016  
DB 9205 TGGAGGTTGCAGTGGGCGGAGATCACAATCAGCTCCAGCTTGGCGGAGCAAGATGA 9264  
  
QY 1017 GACTCTGTCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1066  
DB 9265 AACTCCATCTCAAAAAAATAAAAAAATAAAAAAATAAATAAATAAATAAATAAATAAATA 9314  
  
RESULT 11  
AAL05001  
ID AAL05001 standard; DNA; 12970 BP.

XX AAL05001; 14-SEP-2000; 2000US-0232400.  
AC 14-SEP-2000; 2000US-0232401.  
XX 14-SEP-2000; 2000US-0233053.  
DT 14-SEP-2000; 2000US-0233054.  
XX 14-SEP-2000; 2000US-0233065.  
XX 14-SEP-2000; 2000US-0234223.  
DE 21-SEP-2000; 2000US-0234274.  
XX 21-SEP-2000; 2000US-0234997.  
KW 25-SEP-2000; 2000US-0234998.  
KW 25-SEP-2000; 2000US-0234998.  
XX 26-SEP-2000; 2000US-0234984.  
OS 27-SEP-2000; 2000US-0235834.  
XX 27-SEP-2000; 2000US-0235836.  
PN 29-SEP-2000; 2000US-0236327.  
XX 29-SEP-2000; 2000US-0236367.  
XX 29-SEP-2000; 2000US-0236368.  
XX 29-SEP-2000; 2000US-0236369.  
XX 29-SEP-2000; 2000US-0236370.  
PD 29-SEP-2000; 2000US-0236370.  
XX 02-OCT-2000; 2000US-0236802.  
XX 02-OCT-2000; 2000US-0237037.  
XX 02-OCT-2000; 2000US-0237038.  
XX 02-OCT-2000; 2000US-0237039.  
XX 02-OCT-2000; 2000US-0237040.  
XX 13-OCT-2000; 2000US-0239935.  
XX 13-OCT-2000; 2000US-0239937.  
XX 20-OCT-2000; 2000US-0240960.  
XX 20-OCT-2000; 2000US-0241221.  
XX 20-OCT-2000; 2000US-0241785.  
XX 20-OCT-2000; 2000US-0241786.  
XX 20-OCT-2000; 2000US-0241787.  
XX 20-OCT-2000; 2000US-0241808.  
XX 20-OCT-2000; 2000US-0241809.  
XX 20-OCT-2000; 2000US-0241826.  
XX 01-NOV-2000; 2000US-0244617.  
XX 08-NOV-2000; 2000US-0246474.  
XX 08-NOV-2000; 2000US-0246475.  
XX 08-NOV-2000; 2000US-0246476.  
XX 08-NOV-2000; 2000US-0246477.  
XX 08-NOV-2000; 2000US-0246478.  
XX 08-NOV-2000; 2000US-0246523.  
XX 08-NOV-2000; 2000US-0246524.  
XX 08-NOV-2000; 2000US-0246525.  
XX 08-NOV-2000; 2000US-0246526.  
XX 08-NOV-2000; 2000US-0246527.  
XX 08-NOV-2000; 2000US-0246528.  
XX 08-NOV-2000; 2000US-0246532.  
XX 08-NOV-2000; 2000US-0246609.  
XX 08-NOV-2000; 2000US-0246610.  
XX 08-NOV-2000; 2000US-0246611.  
XX 08-NOV-2000; 2000US-0246613.  
XX 17-NOV-2000; 2000US-0249207.  
XX 17-NOV-2000; 2000US-0249208.  
XX 17-NOV-2000; 2000US-0249209.  
XX 17-NOV-2000; 2000US-0249210.  
XX 17-NOV-2000; 2000US-0249211.  
XX 17-NOV-2000; 2000US-0249212.  
XX 17-NOV-2000; 2000US-0249213.  
XX 17-NOV-2000; 2000US-0249214.  
XX 17-NOV-2000; 2000US-0249215.  
XX 17-NOV-2000; 2000US-0249216.  
XX 17-NOV-2000; 2000US-0249217.  
XX 17-NOV-2000; 2000US-0249218.  
XX 17-NOV-2000; 2000US-0249244.  
XX 17-NOV-2000; 2000US-0249245.  
XX 17-NOV-2000; 2000US-0249264.  
XX 17-NOV-2000; 2000US-0249265.  
XX 17-NOV-2000; 2000US-0249297.  
XX 17-NOV-2000; 2000US-0249299.  
XX 01-DEC-2000; 2000US-0249300.  
XX 01-DEC-2000; 2000US-0250160.  
XX 01-DEC-2000; 2000US-0250391.  
XX 05-DEC-2000; 2000US-0251030.  
XX 05-DEC-2000; 2000US-0251988.  
XX 05-DEC-2000; 2000US-0256719.  
XX 05-DEC-2000; 2000US-0256719.  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225271.  
XX 14-AUG-2000; 2000US-0225275.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226686.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0229287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 01-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 05-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0231242.  
XX 08-SEP-2000; 2000US-0231243.  
XX 08-SEP-2000; 2000US-0231244.  
XX 08-SEP-2000; 2000US-0231413.  
XX 08-SEP-2000; 2000US-0231414.  
XX 08-SEP-2000; 2000US-0232080.  
XX 08-SEP-2000; 2000US-0232081.  
XX 12-SEP-2000; 2000US-0231968.  
XX 14-SEP-2000; 2000US-0232397.  
XX 14-SEP-2000; 2000US-0232398.  
XX 14-SEP-2000; 2000US-0232399.  
XX 14-SEP-2000; 2000US-0232399.  
XX 21-NOV-2001 (first entry)  
Human reproductive system related antigen DNA SEQ ID NO: 7689.  
Human; reproductive system related antigen; reproductive system disorder;  
cancer; gene therapy; ds.  
Homo sapiens.  
W0200155320-A2.  
02-AUG-2001.  
17-JAN-2001; 2001WO-US01339.  
31-JAN-2000; 2000US-0179065.  
04-FEB-2000; 2000US-0180628.  
24-FEB-2000; 2000US-0184664.  
02-MAR-2000; 2000US-0186350.  
16-MAR-2000; 2000US-0189874.  
17-MAR-2000; 2000US-0190076.  
18-APR-2000; 2000US-0198123.  
19-MAY-2000; 2000US-0205515.  
07-JUN-2000; 2000US-0209467.  
28-JUN-2000; 2000US-0214886.  
30-JUN-2000; 2000US-0215135.  
07-JUL-2000; 2000US-0216647.  
07-JUL-2000; 2000US-0216880.  
11-JUL-2000; 2000US-0217487.  
11-JUL-2000; 2000US-0217496.  
14-JUL-2000; 2000US-0218290.  
26-JUL-2000; 2000US-0220963.  
26-JUL-2000; 2000US-0220964.  
14-AUG-2000; 2000US-0224518.  
14-AUG-2000; 2000US-0224519.  
14-AUG-2000; 2000US-0225213.  
14-AUG-2000; 2000US-0225214.  
14-AUG-2000; 2000US-0225266.  
14-AUG-2000; 2000US-0225267.  
14-AUG-2000; 2000US-0225268.  
14-AUG-2000; 2000US-0225270.  
14-AUG-2000; 2000US-0225271.  
14-AUG-2000; 2000US-0225757.  
14-AUG-2000; 2000US-0225758.  
14-AUG-2000; 2000US-0225759.  
18-AUG-2000; 2000US-0226279.  
22-AUG-2000; 2000US-0226681.  
22-AUG-2000; 2000US-0226686.  
22-AUG-2000; 2000US-0227182.  
23-AUG-2000; 2000US-0227009.  
30-AUG-2000; 2000US-0228924.  
01-SEP-2000; 2000US-0229287.  
01-SEP-2000; 2000US-0229343.  
01-SEP-2000; 2000US-0229344.  
01-SEP-2000; 2000US-0229345.  
05-SEP-2000; 2000US-0229509.  
05-SEP-2000; 2000US-0229513.  
06-SEP-2000; 2000US-0230437.  
06-SEP-2000; 2000US-0230438.  
08-SEP-2000; 2000US-0231242.  
08-SEP-2000; 2000US-0231243.  
08-SEP-2000; 2000US-0231244.  
08-SEP-2000; 2000US-0231413.  
08-SEP-2000; 2000US-0231414.  
08-SEP-2000; 2000US-0232080.  
08-SEP-2000; 2000US-0232081.  
12-SEP-2000; 2000US-0231968.  
14-SEP-2000; 2000US-0232397.  
14-SEP-2000; 2000US-0232398.  
14-SEP-2000; 2000US-0232399.





5797	GCACCGCGCTGGCCGCCGCCACATAATTTTCGTTTGAAGTATCTGTACAAATATTTT	5856	PR	14-AUG-2000;	2000US-0225213.
702	CATCAGACTCTACCTTTAAATATGTCAGTTTGGCCAGCCGCTGCTCATGCTGTAA	761	PR	14-AUG-2000;	2000US-0225214.
5857	TA-----CCATTTAAAAACACTGAGTGTGTGGGCACAGTGGCTCACGCTGTAA	5909	PR	14-AUG-2000;	2000US-0225266.
762	TTCCAGCAGCTTTGGGAGGACAGAGTGGGTGAATCACTTGAGCCAGAGTTTGAGACCA	821	PR	14-AUG-2000;	2000US-0225268.
5910	TCCCAGAACTTTGGGAGGCCGAGCGGACGACATCAA--GAGGTCAGGAGATGGAGACCA	5967	PR	14-AUG-2000;	2000US-0225270.
822	CCTGGCCAACTGTGTGAATCTGTCTTTACTAAAAAATAAATAAATAAATAAATAA	881	PR	14-AUG-2000;	2000US-0225271.
5968	CCTGGCTAAACAGGTGAACCTCTGTCTACT-----AAAAATAAATAAATAA	6015	PR	14-AUG-2000;	2000US-0225272.
882	GCCGGGTGCTGGGGCACACCTCTGAATCCAGCATCTGTCGAGGCTGAGGACGAGAGT	941	PR	14-AUG-2000;	2000US-0225273.
6016	GCCGGGCATGTTGGCAGCGCCTGTAGTCCAGCTACTCGGGAGGCTGAGCGAGGAGAT	6075	PR	14-AUG-2000;	2000US-0225274.
942	CAGTTGAACCTGGAGGCGGAGGTTGCAGTGGCGGAGATCACATCACCGCCCTCCAGCC	1001	PR	14-AUG-2000;	2000US-0225275.
6076	GGCGTGAACCCAGGAAGTGGAGCTTGCAGTGAGCGGAGATCACGCCACTGCACTCCAGCC	6135	PR	14-AUG-2000;	2000US-0225276.
1002	TGGCGGCACAGACGAAGACTCTGTCTCAATAAATAAATAAATAAATAAATAAATAA	1061	PR	14-AUG-2000;	2000US-0225277.
6136	TGGCGGACAGTSCAAGACTCCGCTCTCAAAAAATAAATAAATAAATAAATAAATAA	6195	PR	14-AUG-2000;	2000US-0225278.
1062	TTGTACCTTAGTTATATCTAATAAAA	1086	PR	14-AUG-2000;	2000US-0225279.
6196	TTGAGTTGTAGGAATGCTTTATATA	6220	PR	14-AUG-2000;	2000US-0225280.
RESULT 13					
AAK89986/C					
ID	AAK89986 standard; DNA; 32224 BP.				
XX	AAK89986;				
AC	(first entry)				
DT	05-NOV-2001				
DE	Human digestive system antigen genomic sequence SEQ ID NO: 3562.				
DE	Human; digestive system antigen; gene therapy; cancer; appendicitis;				
KW	ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;				
KW	digestive system disorder; Meckel's diverticulum; ds.				
OS	Homo sapiens.				
OS	WO200155314-A2.				
PN	02-AUG-2001.				
PN	17-JAN-2001; 2001WO-US01324.				
XX	31-JAN-2000; 2000US-0179065.				
XX	04-FEB-2000; 2000US-0180628.				
XX	24-FEB-2000; 2000US-0184664.				
XX	02-MAR-2000; 2000US-0186350.				
XX	16-MAR-2000; 2000US-0189874.				
XX	17-MAR-2000; 2000US-0190076.				
XX	18-APR-2000; 2000US-0198123.				
XX	19-MAY-2000; 2000US-0205515.				
XX	07-JUN-2000; 2000US-0209467.				
XX	28-JUN-2000; 2000US-0214886.				
XX	30-JUN-2000; 2000US-0215135.				
XX	07-JUL-2000; 2000US-0216647.				
XX	07-JUL-2000; 2000US-0216880.				
XX	11-JUL-2000; 2000US-0217487.				
XX	11-JUL-2000; 2000US-0217496.				
XX	14-JUL-2000; 2000US-0218290.				
XX	26-JUL-2000; 2000US-0220963.				
XX	26-JUL-2000; 2000US-0220964.				
XX	14-AUG-2000; 2000US-0224518.				
XX	14-AUG-2000; 2000US-0224519.				



PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 06-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-541565/60.  
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating nervous system  
XX cancers and metastases -  
XX









GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 5, 2003, 02:24:45 ; Search time 3765.66 Seconds

(without alignments)

12908.483 Million cell updates/sec

Title: US-09-936-271B-13\_COPY\_1\_2000

Perfect score: 2000

Sequence: 1 gggccagagtgaaggcaag.....gtgcgtctgcaccacatc 2000

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_nam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	397.6	19.9	2371	28	AF101960 AF101960
2	373.8	18.7	2971	28	AF101960 AF101960
3	352.2	17.6	929	9	AL580585 AL580585
C 4	337.4	16.9	652	9	AU120416 AU120416

C 5	337.2	16.9	1641	11	BC033224	BC033224 Homo sapi
C 6	336	16.8	2772	11	BC038630	BC038630 Homo sapi
C 7	334.4	16.7	698	12	BM999040	BM999040 UI-H-D10-
8	332	16.6	969	13	BX415714	BX415714 BX415714
C 9	329.6	16.5	929	9	AL580585	AL580585 AL580585
C 10	325.2	16.3	652	9	AU120416	AU120416 AU120416
C 11	324	16.2	701	10	BE744242	BE744242 601576428
C 12	321.6	16.1	2772	11	BC038630	BC038630 Homo sapi
C 13	316.8	15.8	381	29	CC061638	CC061638 MUGQ_CH25
C 14	315.6	15.8	709	29	AG010128	AG010128 Homo sapi
15	315	15.8	710	14	CA428305	CA428305 UI-H-DP0-
16	312.6	15.6	542	9	AI679782	AI679782 tu7610.x
17	312.2	15.6	495	9	AI963720	AI963720 wf6604.x
C 18	311.4	15.6	3491	11	BC039100	BC039100 Homo sapi
C 19	311.2	15.6	925	9	AL558859	AL558859 AL558859
C 20	309.6	15.5	1910	11	BC035771	BC035771 Homo sapi
C 21	309.4	15.5	3491	11	BC039100	BC039100 Homo sapi
C 22	308.4	15.4	764	13	BU617735	BU617735 UI-H-DFO-
23	305.4	15.3	2076	11	BC035510	BC035510 Homo sapi
24	305	15.2	834	29	BZ612098	BZ612098 WHADP72TR
C 25	302.8	15.1	322	13	BU941317	BU941317 AGENCOURT
C 26	301.6	15.1	398	9	AL596113	AL596113 DREZP686A
C 27	301.2	15.1	868	10	BF337291	BF337291 602034658
C 28	299.8	15.0	1073	29	BZ601227	BZ601227 WHADC89TR
29	299.8	15.0	1910	11	BC035771	BC035771 Homo sapi
C 30	299.8	15.0	2076	11	BC035510	BC035510 Homo sapi
C 31	299.6	15.0	829	29	BZ609393	BZ609393 WHACI48TF
C 32	299.4	15.0	599	13	BU729004	BU729004 UI-E-CQ1-
C 33	299.2	15.0	691	9	AL043009	AL043009 DREZP4340
C 34	298.8	14.9	584	12	BM701794	BM701794 UI-E-CQ1-
C 35	298.6	14.9	698	12	BM999040	BM999040 UI-H-DFO-
C 36	297.6	14.9	844	28	AQ748733	AQ748733 HS_5540.A
37	297.2	14.9	755	13	BX352666	BX352666 BX352666
38	297	14.8	957	9	AL573566	AL573566 AL573566
C 39	296	14.8	400	2	HSW072571	Bx482403 Homo sapi
C 40	296	14.8	910	13	BQ722917	BQ722917 AGENCOURT
C 41	295.8	14.8	540	13	BU617996	BU617996 UI-H-DFO-
C 42	295	14.8	564	10	BG529995	BG529995 602558836
43	294.8	14.7	815	29	BZ608327	BZ608327 WHACC28TR
C 44	294.6	14.7	533	28	AQ082371	AQ082371 RPC111-55
45	294.4	14.7	2284	11	BC036238	BC036238 Homo sapi

## ALIGNMENTS

RESULT 1  
AF101960/c

LOCUS

DEFINITION AF101960 Human Homo sapiens genomic clone pTWB59.14, genomic survey

sequence.

ACCESSION AF101960

VERSION AF101960.1 GI:4193786

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2971)

AUTHORS Bepler,G., O'Briant,K.C., Kim,Y.C., Schreiber,G. and Pittorle,D.M.

TITLE A 1.4-Mb high-resolution physical map and contig of chromosome segment 11p15.5 and genes in the LOH11A metastasis suppressor region

JOURNAL Genomics 55 (2), 164-175 (1999)

MEDLINE 99134294

PUBMED 9933563

COMMENT Contact: Bepler G

Duke University Medical Center

Box 2610, MSRB, Room 117, Durham, NC 27710, USA

part of a 1.4 megabase contig including the LOH11A metastasis suppressor region Bin T

Class: unknown.

2971 bp DNA linear GSS 06-NOV-2000

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.





```

source      1. .652
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEM1000637"
/tissue_type="whole embryo, mainly body"
/dev_stage="embryo, 10 weeks"
/clone_lib="HEM1000637"
/notes="Vector: pME18SFL3"
BASE COUNT 155 a 161 c 163 g 167 t. 6 others
ORIGIN

Query Match 16.9%; Score 337.4; DB 9; Length 652;
Best Local Similarity 75.7%; Pred. No. 6.9e+02;
Matches 494; Conservative 0; Mismatches 125; Indels 34; Gaps 5;

Qy 400 CAGCTGGAGTGCAGTGGCGGATCTCGGCTCACTGCAAGCTCCGCTCCCGGTTTACG 459
Db 642 CAGTGTGAGTGCAGTGGCGGATCTCGGCTCACTGCAAGCTCCGCTCCCGGTTTACG 583
Qy 460 CCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 519
Db 582 CGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523
Qy 520 GCTA-----ATTTTTTTTATTTTATTTAGTAGACGGGTTTTCACGTTTATGCGGG 571
Db 522 GCTATCTTTTNTTTTTTTTTTTTTTTTAAAGTAGACAGAGGTTTCACTGTGTAGCCAGG 463
Qy 572 ATGGCTCGATCTCTGACCTCGATCCGCGCGCTCGGCTCCCAAGTGTGGGATT 631
Db 462 ATGGCTCGATCTCTGATCTCGTGTGATCTGCGCCACCTCGGCTTCCCAAGTGTGGGATT 403
Qy 632 ACAGCGCTGAGCCCGCGCGCGCC--ATGATCATCTTCTTGACTATGCTGATGTGACA 689
Db 402 ACAGCGCTGAGCCCGCGCGCGCGCTGATATATATATTTTGTGATTTTAAAGGCC 343
Qy 690 AGTACCTAAAGCCATGACATCTACCTTTAAATATGATGATTTGGCGCAGCAGCGTGGC 749
Db 342 AAGTCAGAAATCAACACACACACACAAATAAA-----GGCCAGCAGCGTGGC 294
Qy 750 TCATGCTCTGTAATTCAGACACTTTGGGAGGAGAGGTGGTGAATCACTTTCAGGCGCAGGA 809
Db 293 TCAGCGCTGTAATTCAGACACTTCGGAGGCGGAGGAGGTGGTGAATCACTTTCAGGCGCAGGA 234
Qy 810 GTTTGAGACACCGCTGGCCACATGTTGAACTCTGTTTACTTAAATAAATAAATAAATAA 869
Db 233 GTTTGAGACACCGCTGGCCACATGTTGAACTCTGTTTACTTAAATAAATAAATAAATAA 186
Qy 870 AAAAAAAATCAGCGGGGTGCTGGGCGCACACCTGTAATCCAGCTATGCTGGAGGCTG 929
Db 185 ATACAAAAATTAGTGGGCGTGTGATGTGAGTGGCTGTGTAATCCAGCTACTCTGGAGGCTG 126
Qy 930 AGGCAGAGAGTCACTTGAACCTTGGAGGCGAGGTTGCACTGGCGGAGATACATCATCAC 989
Db 125 AGGTAGAGAAATGCTTGAACCCAGAGGAGTGGAGGTTGCACGAGCGAGATCCGCGCAT 66
Qy 990 CGCCTCCAGCTCGGCGAC--AGAGCAAGACTCTGCTCAATAAATAAATAAATAA 1041
Db 65 TGCACCTCAGGCTGGAGCGCAAGAGCAAACTCTGTCTCAAAAAATAAATAAATAA 13

RESULT 5
BC033224/c
LOCUS BC033224 1641 bp mRNA linear HTC 27-JUN-2002
DEFINITION Homo sapiens, clone IMAGE:5022014, mRNA.
ACCESSION BC033224
VERSION BC033224.1 GI:21619936
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1641)

```

```

AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabps-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 43 Row: j Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: frame shifted.
FEATURES
Location/Qualifiers
source 1. .1641
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5022014"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH_MGC_17"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
BASE COUNT 468 a 364 c 447 g 362 t
ORIGIN
Query Match 16.9%; Score 337.2; DB 11; Length 1641;
Best Local Similarity 75.3%; Pred. No. 3.4e+02;
Matches 500; Conservative 0; Mismatches 128; Indels 36; Gaps 5;
Qy 348 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 407
Db 1118 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1059
Qy 408 AGTCAGTGGCGGATCTCGGCTCACTGCAAGCTCCGCGTCCCGGTTTACGCCATCTC 467
Db 1058 AGTCAGTGGCGGATCTCGGCTCACTGCAAGCTCTGCTCTCTGTTTACGCCATCTC 999
Qy 468 CTGCTCAGCTCCCAAGTAGCTGGGACTACAGCGCGCGGCTACGCGCGGCTAAATTT 527
Db 998 CTGCTCAGCTCCCAAGTAGCTGGGACTACAGCGCGCGGCTACGCGCGGCTAAATTT 940
Qy 528 TTTTGTATTTTATTTAGAGCGGGTTTACCGTTTATAGCGGGATGGCTCGATCTCCT 587
Db 939 TTTTGTATTTTATTTAGAGCGGGTTTACCGTTTATAGCGGGATGGCTCGATCTCCT 880
Qy 588 GACCTCGTATCCCGCGCTCGGCTCCCAAGTGTGGGATTTACAGCGGTGAGCCACC 647
Db 879 GACCTCGTATCCCGCGCTCGGCTCCCAAGTGTGGGATTTACAGCGGTGAGCCACC 820
Qy 648 GCGCCCGGCGCATGATCATCTTCTTTGACTATGCTGATGTGACAAGTACCTAAAGCCATCAG 707
Db 819 ATGCCCGGCTAAACTCTCATTTCTAAATTTGGACAATTTTAAAGAAGTATTTTACTCCACA- 761
Qy 708 ACTCTACCTTTAAATATGACAGTTTGGCCAGGACCGTGGCTCATGCTTAATTCAG 767

```













Dn 201 GCCACGAGCCCGCAGAGGCAGTGGCCAGGAAGGTGCAGGCCCTGAGAATTCGCCGCGTG 142

Oy 1889 AGCTGGGAGCAAAATCCCCACCCTACTCTGGGGGACAGGCAGTAAGTAGACCTGTGTGAGG 1948  
| | | | |  
Dn 141 AACTGGNAAGGAATCCCACACCCCTACCTTTGGGGGACAGGCAGGTGAGACCTTGGGGAGG 82  
Oy 1949 GTGGCTCAGCAGCAGGGAAGGAGAGGTGCTGTGGGTCTGTGACCCCATC 2000  
| | | | |  
Dn 81 GTGGCTCAGCAGCAGGGAAGGAGAGGTGCTGTGGGTCTGTGACCCCATC 30

RESULT 14  
AG010128  
LOCUS  
DEFINITION Homo sapiens genomic DNA, 21q region, clone: f2g2x4, genomic survey sequence.  
ACCESSION AG010128 AG003939 linear GSS 14-APR-1999  
VERSION AG010128.1 GI:3294404  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 709)  
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.  
TITLE Homo sapiens genomic DNA, chromosome 21q  
JOURNAL Published Only in DataBase (1998)  
REFERENCE 2 (bases 1 to 709)  
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (08-JUL-1998) Masahira Hattori, Kitasato University,  
Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1,  
Sagamihara 228, Japan (E-mail:hattoriehg.cims.u-tokyo.ac.jp,  
Tel.:0427-78-9732, Fax:0427-78-9561)  
COMMENT On Feb 5, 1999 this sequence version replaced gi:2760798.  
AG003939: Submitted (09-Jan-1998).  
  
FEATURES source Location/Qualifiers  
1..709  
    /organism="Homo sapiens"  
    /mol\_type="genomic DNA"  
    /db\_xref="taxon:9606"  
    /chromosome="21"  
    /map="21q"  
    /cloned="f2g2x4"

BASE COUNT 171 a 179 c 184 g 168 t 7 others

ORIGIN  
  
Query Match 15.8%; Score 315.6; DB 29; Length 709;  
Best Local Similarity 73.3%; Pred. No. 1.4e+03;  
Matches 499; Conservative 0; Mismatches 139; Indels 43; Gaps 6;

Oy 375 TGACAGCGAATCTCGCTCTCTCGCCCAAGGCTGAGTGCAAGTCTCGGGATCTCGGCTCAC 434  
| | | | |  
Dn 7 TGAATTCCGNCCCGGGAGGTCGCCCAAGCTGGAGTGCAAGTCTCGGGATCTCGGCTCAC 66  
| | | | |  
Oy 435 CCAAGCTCCGCCTCCCGGGTTCACGCCATTCTCTGCCTCAGCCTCCCAAGTAGCTGGGA 494  
| | | | |  
Dn 67 GCRAAGCTCCGCTCCCGGGTTCACGCCATTCTCTGCCTCAGCCTCCCAAGTAGCTGGGA 126  
| | | | |  
Oy 495 CTACAGCGCCCGCCACTACGCCCGGCTAATTTTTTGTATTTTGTAGTAGACGGGGTT 554  
| | | | |  
Dn 127 CTCAGCGCCCGCCACTACGCCCGGCTAATTTTTTGTATTTTGTAGTAGACGGGGTT 186  
| | | | |  
Oy 555 TCACCGTTTTCAGCGGGATCGCTGATCTCTGACCTCGTGATCGCCCGCTCGGCT 614  
| | | | |  
Dn 187 TCACCATTTTTCAGCGGGATCGCTGATCTCTGACCTCGTGATCGCCCGCTCGGCT 246  
| | | | |  
Oy 615 CCACAAGTGTGGATTTACAGCGGTGAGCCACCGCGCGCCGATGATCATCTCTTGAC 674  
| | | | |  
Dn 247 CCCAAGTGTGGATTTACAGCGGTGAG-CACCGCGCGCGCTGAATTTGTTGTTG 305  
| | | | |  
Oy 675 TATGCTGATGTGACAAGTACCTAAAGCCATCAGACTCTACCCCTTTAATAATGCAAGTTTG 734  
| | | | |



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2003, 20:39:54 ; Search time 511.264 Seconds

(without alignments)

10934.739 Million cell updates/sec

Title: US-09-936-271B-13\_COPY\_9500\_11570

Perfect score: 2071

Sequence: 1 aaacagacacaaaacttctct.....ccagaaataaactgagaag 2071

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

Result No.	Score	Query Match	Length	DB ID	Description
1	2063	99.6	11570	21	AAA95905 Human KLK-L2 gene.
C 2	482	23.3	586	24	ABL63579 Breast cancer rela
C 3	482	23.3	586	24	ABL63996 Breast cancer rela
C 4	482	23.3	735	21	RAC79469 cDNA sequence of h
C 5	482	23.3	735	24	ABK29013 Human breast tumou
C 6	482	23.3	738	20	AAH84240 DNA encoding huma
C 7	482	23.3	1302	20	AAH57989 Human BS247 specif
C 8	482	23.3	1302	20	AAH57990 Human BS247 specif

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

9	482	23.3	1381	20	AAZ06259 Human secreted pro
10	482	23.3	1439	20	AAZ06260 Human secreted pro
11	482	23.3	1499	20	AX60578 Human keratinocyte
12	482	23.3	1516	20	AAZ06245 Human secreted pro
13	482	23.3	1570	21	AAZ65070 Membrane-bound pro
14	482	23.3	1570	22	AAZ21471 Human cDNA sequenc
15	482	23.3	1570	22	AAZ44216 Human PRO1132 (UNQ
16	482	23.3	1570	24	ABK28605 Human DNA59767-148
17	482	23.3	1570	25	ACA03830 cDNA encoding huma
18	482	23.3	1570	25	ACA04251 Human cDNA encodin
19	482	23.3	1570	25	ABX89368 DNA encoding novel
20	482	23.3	1570	25	ABX80318 Human secreted or
21	482	23.3	1570	25	ABX80822 Human secreted/tr
22	482	23.3	1570	25	ABX81205 Human secreted or
23	482	23.3	1570	25	ABX90295 Human secreted/tr
24	482	23.3	1570	25	ABX77906 Human PRO polynucl
25	482	23.3	1570	25	ABX79502 Human secreted/tr
26	482	23.3	1570	25	ABX64141 cDNA encoding huma
27	482	23.3	1570	25	ABX17105 Human PRO polynucl
28	481	23.2	1539	20	AAZ23317 Human PDSF-1 DNA.
29	476	23.0	1476	20	AAK16295 Human kallikrein e
30	460	22.2	1504	24	ABK30231 Human G-protein-co
31	449	21.7	673	22	AAF44901 Human breast cance
32	259	12.5	259	20	AAH57987 Human BS247 specif
33	251	12.1	254	20	AAH57988 Human BS247 specif
34	243.2	11.7	130263	24	ABK83573 Human cDNA differe
C 35	226	10.9	757	22	AAI96793 Human neuroblastom
C 36	225.4	10.9	240825	22	AAF24497 Human PG-3 gene.
C 37	225.4	10.9	240825	24	ABQ81802 Human PG-3 gene SE
C 38	222.6	10.7	11084	20	AAZ22302 Human IL-1ra BAC c
C 39	219.4	10.6	70000	24	AAZ42934 Human phospholipas
C 40	219.2	10.6	977	22	AAK82079 Human immune/haema
C 41	219.2	10.6	3830	20	AAH02984 Human IL-1ra BAC c
C 42	219.2	10.6	503	23	ABV54065 Human prostate exp
C 43	218.6	10.6	25464	19	AAH57274 Human flavin-conta
C 44	218.4	10.5	1446	22	AAH99110 Human EST-derived
C 45	218	10.5	4824	24	ABK70289 Human lung cancer

#### ALIGNMENTS

RESULT 1	
AAA95905	
ID AAA95905 standard; DNA; 11570 BP.	
XX AAA95905;	
AC AAA95905;	
DT 02-FEB-2001 (first entry)	
XX Human KLK-L2 gene.	
DE	
XX Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;	
KW Kallikrein-like protein; serine protease; cytosolic; cancer;	
KW prostrate cancer; ds.	
XX	
OS Homo sapiens.	
XX	
PN WO200053776-A2.	
PD	
XX 14-SEP-2000.	
XX	
PF 09-MAR-2000; 2000WO-CA00258.	
XX	
PR 11-MAR-1999; 99US-0124260.	
PR 01-APR-1999; 99US-0127386.	
PR 21-JUL-1999; 99US-0144919.	
XX	
PA (MOUN ) MOUNT SINAI HOSPITAL.	
XX	
PI Yousef GM, Diamandis EP;	
XX	
DR WPI; 2000-587440/55.	

DR P-PSDB: AAB21296.

XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L  
PT protein mediated disorders, especially cancer.

XX Claim 1: Page 143-149; 184pp; English.

XX The present sequence is the coding sequence of the human KLK-L2 gene,  
CC which encodes a kallikrein-like protein. Kallikreins and kallikrein-like  
CC proteins are a subgroup of the serine protease enzyme family. They  
CC catalyze the selective cleavage of specific polypeptide precursors to  
CC release peptides with potent biological activity. Nucleic acids encoding  
CC kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and  
CC KLK-L6 have been isolated. The proteins are useful in the treatment,  
CC monitoring and diagnosis of cancers, especially prostate cancer. They  
CC can also be used to identify a substance that can associate with or  
CC mediate the biological activity of the proteins. Antibodies can be used  
CC to treat conditions mediated by the kallikrein-like proteins.

XX Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 other;

Query Match 99.68; Score 2063; DB 21; Length 11570;  
Best Local Similarity 99.88; Pred. No. 0;  
Matches 2066; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AAACAGACCAAAACCTTCCTGCCCGCTGGACCTCATGTTCCCAAGTGGAGACAGCAA 60  
Db 9500 AAACAGACCAAAACCTTCCTGCCCGCTGGACCTCATGTTCCCAAGTGGAGACAGCAA 9559  
Qy 61 TAAAGAGATAGATAAATATAGTAATTAATAAATAAATAAATAAATAGCCGGGTGGTGG 120  
Db 9560 TAAAGAGATAGATAAATATAGTAATTAATAAATAAATAAATAAATAGCCGGGTGGTGG 9619  
Qy 121 CTTGCACCTGTAGTTCAGCTTCTGGAGGCTGAGTGGAGAAATGCTTGGACCCAAA 180  
Db 9620 CTTGCACCTGTAGTTCAGCTTCTGGAGGCTGAGTGGAGAAATGCTTGGACCCAAA 9679  
Qy 181 CGTTTGGAGCTCGGTAAAGCCATGACTGCTGCTGCTCCAGACAGACCTCGGTGTA 240  
Db 9680 CGTTTGGAGCTCGGTAAAGCCATGACTGCTGCTGCTCCAGACAGACCTCGGTGTA 9739  
Qy 241 CAAGCAGACGCTTTTGTGAGAAAGAAATAAATAAATAAATAAAGAGAGAGAGAG 300  
Db 9740 CAAGCAGACGCTTTTGTGAGAAAGAAATAAATAAATAAATAAAGAGAGAGAGAG 9799  
Qy 301 AAAG 360  
Db 9800 AAAG 9859  
Qy 361 GAAAG 420  
Db 9860 GAAAG 9919  
Qy 421 GAGNAGAT 480  
Db 9920 GAGNAGAT 9979  
Qy 481 TATTATAGAGGTAATATATAGGAGGTATGCGGAATTTGAAGACAGAGAAACAAATTAG 540  
Db 9980 TATTATAGAGGTAATATATAGGAGGTATGCGGAATTTGAAGACAGAGAAACAAATTAG 10039  
Qy 541 TCCAGCGAATGGATTTCTATTGGAGGTATTTGCCCTTAGAAGACACTGGCAATACCA 600  
Db 10040 TCCAGCGAATGGATTTCTATTGGAGGTATTTGCCCTTAGAAGACACTGGCAATACCA 10099  
Qy 601 GGAGACATTTTGGTGTCAACACTATATGAGGGGCAATTAAGTGGCAATAGATAGA 660  
Db 10100 GGAGACATTTTGGTGTCAACACTATATGAGGGGCAATTAAGTGGCAATAGATAGA 10159  
Qy 661 TGCCAAAGTGTGCTGTCAACATGCTATGATCAGACAGGCGCTCCCAACAAACCAATT 720  
Db 10160 TGCCAAAGTGTGCTGTCAACATGCTATGATCAGACAGGCGCTCCCAACAAACCAATT 10219

Qy 721 ATCCAGCTTCAGATGCCACAGTCCAGATCCAGAGAAACCTCATCCAGGGGTGAGAAC 780  
Db 10220 ATCCAGCTTCAGATGCCACAGTCCAGATCCAGAGAAACCTCATCCAGGGGTGAGAAC 10279  
Qy 781 CGTATTTTTCAGAGAGAGGTATAGGATGGGTGGTGGAGAAATGGGGAAGAGGTGT 840  
Db 10280 CGTATTTTTCAGAGAGAGGTATAGGATGGGTGGTGGAGAAATGGGGAAGAGGTGT 10339  
Qy 841 GTGTCCAGTAAAGAGAAATAGGCTGCACAGGCTGGAGGGGAGAGTCCAGAGAAAGGGA 900  
Db 10340 GTGTCCAGTAAAGAGAAATAGGCTGCACAGGCTGGAGGGGAGAGTCCAGAGAAAGGGA 10399  
Qy 901 GCGCGAGAGATACACAGTATGAGGAGACAGGCTGGACAGAAAGTAGAGACAGATTCGA 960  
Db 10400 GCGCGAGAGATACACAGTATGAGGAGACAGGCTGGACAGAAAGTAGAGACAGATTCGA 10459  
Qy 961 GATGTGAGAGAGAGGTTCACAGACCCCGGAAATGATGTGGACAAACAGGAATCTGG 1020  
Db 10460 GATGTGAGAGAGAGGTTCACAGACCCCGGAAATGATGTGGACAAACAGGAATCTGG 10519  
Qy 1021 AAGAGGAAGATGAGTGGAGAGTGCACAAATGGGCTCTAAAGGTTGAACCTTGGAGGCCAGG 1080  
Db 10520 AAGAGGAAGATGAGTGGAGAGTGCACAAATGGGCTCTAAAGGTTGAACCTTGGAGGCCAGG 10579  
Qy 1081 CATGGTGGCTCACGCCCTGTAAATCCCAACACTTTGGAGGCTGAGTGGGCGGAATCAGTTGA 1140  
Db 10580 CATGGTGGCTCACGCCCTGTAAATCCCAACACTTTGGAGGCTGAGTGGGCGGAATCAGTTGA 10639  
Qy 1141 GCGCAGAGATTGAGAGACAGCCTGGCCAAACATGTTGAAACCCCTCTCTACAAAAAAT 1200  
Db 10640 GCGCAGAGATTGAGAGACAGCCTGGCCAAACATGTTGAAACCCCTCTCTACAAAAAAT 10599  
Qy 1201 ACAAAAAATTAGCCGGGTGGTGGATGGACACACTGTAGTACACAGTACTTTGGAGGCTGA 1260  
Db 10700 ACAAAAAATTAGCCGGGTGGTGGATGGACACACTGTAGTACACAGTACTTTGGAGGCTGA 10759  
Qy 1261 GCGAGAGAAATTTGTTGAAACCCGGAGATGGAGGCTGCAGTACAGTGCAGGCTCAGT 1320  
Db 10760 GCGAGAGAAATTTGTTGAAACCCGGAGATGGAGGCTGCAGTACAGTGCAGGCTCAGT 10819  
Qy 1321 GCGCTCAACCTGGGCAACAGAGTAAGACTCCATCTCAAAAAAAGGCTGATTTGG 1380  
Db 10820 GCGCTCAACCTGGGCAACAGAGTAAGACTCCATCTCAAAAAAAGGCTGATTTGG 10879  
Qy 1381 AGTGAATATTAATAACATTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440  
Db 10880 AGTGAATATTAATAACATTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10939  
Qy 1441 TTTTCTGCATTTCTTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
Db 10940 TTTTCTGCATTTCTTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10999  
Qy 1501 TCCATCTATGGGCTCTCTGGTCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560  
Db 11000 TCCATCTATGGGCTCTCTGGTCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11059  
Qy 1561 GCCTCTCTCATGCCCT 1620  
Db 11060 GCCTCTCTCATGCCCT 11119  
Qy 1621 ATGCTCTCTCAGGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
Db 11120 ATGCTCTCTCAGGGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 11179  
Qy 1681 CGGGTGTCTACACGAACCTCTGCAAGTTACCAAGTGGATCCAGAAACCAATCCAGGCA 1740  
Db 11180 CGGGTGTCTACACGAACCTCTGCAAGTTACCAAGTGGATCCAGAAACCAATCCAGGCA 11239  
Qy 1741 ACTCCTGAGTATCCAGGACTCAGCACACCGGATCCCACTGCTGCTGAGGAGAGGCC 1800  
Db 11240 ACTCCTGAGTATCCAGGACTCAGCACACCGGATCCCACTGCTGCTGAGGAGAGGCC 11299  
Qy 1801 TGACACTCTCTTTCAGACCCCTCATTTCTCCAGAGATGTTGAGAAATGTTCTCTCTCCAG 1860



Db	11300	TGACACTCCTTTTCAGACCCTCAATTCCTCCACAGATGTTGAGAATGTCATCTCTCCAG	11359
QY	1861	CGCTTGACCCCATATGCTCTCTGGACTCAGGGTCTGCTTCCCCACATTTGGGTGACCCGTG	1920
Db	11360	CGCCTGACCCCATGCTCTCTGGACTCAGGGTCTGCTTCCCCACATTTGGGTGACCCGTG	11419
QY	1921	CTCTCTAGTTGAACCTCTGGGAACAATTTCCAAAACGTGCCAGGGGGGGGTTGGCTCTCA	1980
Db	11420	CTCTCTAGTTGAACCTCTGGGAACAATTTCCAAAACGTGCCAGGGGGGGGTTGGCTCTCA	11479
QY	1981	ATCTCCCTGGGGCACTTTTCATCTCTCAAGCTCAGGGCCCATCCCTCTCTCTGAGCTCTGAC	2040
Db	11480	ATCTCCCTGGGGCACTTTTCATCTCTCAAGCTCAGGGCCCATCCCTCTCTCTGAGCTCTGAC	11539
QY	2041	CCAAATTTAGTCCCGAATAAATCGAAG	2071
Db	11540	CCAAATTTAGTCCCGAATAAATCGAAG	11570

RESULT 2	
ABL63579/c	
ID	ABL63579 standard; DNA; 586 BP.
XX	
XX	ABL63579;
XX	
XX	
DT	15-MAY-2002 (first entry)
XX	
XX	Breast cancer related gene sequence SEQ ID NO:1916.
DE	
XX	
XX	
KW	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid; stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous; cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW	
KW	

DT	15-MAY-2002	(first entry)
DE	Breast cancer related gene sequence SEQ ID NO:1916.	
XX	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;	
XX	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;	
KW	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;	
KW	gene; ds.	
XX		
XX	Homo sapiens.	
OS		
XX		
XX	WO200194629-A2.	
PN		
XX	13-DEC-2001.	
PD		

30-MAY-2001;	2001WO-US10838.
05-JUN-2000;	2000US-209473P.
05-JUN-2000;	2000US-209531P.
18-SEP-2000;	2000US-233133P.
18-SEP-2000;	2000US-233617P.
20-SEP-2000;	2000US-234009P.
20-SEP-2000;	2000US-234034P.
20-SEP-2000;	2000US-234052P.
22-SEP-2000;	2000US-234509P.
22-SEP-2000;	2000US-234567P.
25-SEP-2000;	2000US-234923P.
25-SEP-2000;	2000US-234924P.
25-SEP-2000;	2000US-235077P.
25-SEP-2000;	2000US-235082P.
25-SEP-2000;	2000US-235134P.
25-SEP-2000;	2000US-235280P.
26-SEP-2000;	2000US-235637P.
26-SEP-2000;	2000US-235638P.
27-SEP-2000;	2000US-235711P.
27-SEP-2000;	2000US-235720P.
27-SEP-2000;	2000US-235840P.
27-SEP-2000;	2000US-235863P.
28-SEP-2000;	2000US-236028P.
28-SEP-2000;	2000US-236032P.
28-SEP-2000;	2000US-236033P.
28-SEP-2000;	2000US-236034P.
28-SEP-2000;	2000US-236109P.
28-SEP-2000;	2000US-236111P.
29-SEP-2000;	2000US-236842P.
29-SEP-2000;	2000US-236891P.
02-OCT-2000;	2000US-237172P.

PR	02-OCT-2000;	2000US-237173P.
PR	02-OCT-2000;	2000US-237278P.
PR	02-OCT-2000;	2000US-237294P.
PR	02-OCT-2000;	2000US-237295P.
PR	02-OCT-2000;	2000US-237316P.
PR	03-OCT-2000;	2000US-237425P.
PR	03-OCT-2000;	2000US-237598P.
PR	03-OCT-2000;	2000US-237604P.
PR	03-OCT-2000;	2000US-237606P.
PR	03-OCT-2000;	2000US-237608P.
PR	01-NOV-2000;	2000US-244867P.
PR	01-NOV-2000;	2000US-245084P.
XX	(AVAL - )	AVALON PHARM.
XX	PA	PA
XX	Young PE,	Augustus M,
PI	Soppet DR,	Weaver Z;
PI		
XX		
DR	WPI:	2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a  
chemical agent to be tested for anti-neoplastic activity, and  
determining a change in expression of a gene of a signature gene set -  
Claim 1; SEQ ID 1916; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's tumour.

Sequence 586 BP; 130 A; 139 C; 179 G; 136 T; 2 other;

	Query Match	23.3%;	Score 482;	DB 24;	Length 586;	
	Best Local Similarity	100.0%;	Prod. No. 5.5e-81;			
	Matches 482;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1590	CAGGGTGATTCTCGGGGGGCGCTGTGGTCTCAATGGCTCCCTGCAGGAGCTCGTGTCCTGG	1649			
Db	501	CAGGGTGATTCTCGGGGGGCGCTGTGGTCTCAATGGCTCCCTGCAGGAGCTCGTGTCCTGG	442			
Qy	1650	GGAGATTACCTTGTGCGCGGGCCCAACAGACCGGGGTGCTACAGCAACCTCTCGCAAGTTC	1709			
Db	441	GGAGATTACCTTGTGCGCGGGCCCAACAGACCGGGGTGCTACAGCAACCTCTCGCAAGTTC	382			
Qy	1710	ACCAAGTGGATCCAGGAACCAATCCAGGGCAACTCTGTAGTCAATCCAGAGCTCAGCAC	1769			
Db	381	ACCAAGTGGATCCAGGAACCAATCCAGGGCAACTCTGTAGTCAATCCAGAGCTCAGCAC	322			
Qy	1770	CCGGCATCCCACTGCTGCAGGAGAGCCCTGACACTCCTTTTCAGACCCCTCAATTCCTTC	1829			
Db	321	CCGGCATCCCACTGCTGCAGGAGAGCCCTGACACTCCTTTTCAGACCCCTCAATTCCTTC	262			
Qy	1830	CCAGAGATGTTGAGAAATGTTCACTCTCCAGGCCCTGACCCCACTGTCTCTGTGACATCAGG	1889			
Db	261	CCAGAGATGTTGAGAAATGTTCACTCTCCAGGCCCTGACCCCACTGTCTCTGTGACATCAGG	202			
Qy	1890	GTCTGCTTCCCCACATATGGGCTGACCGTGTCTCTCTAGTGTGAACCTGGGAACAATTC	1949			
Db	201	GTCTGCTTCCCCACATATGGGCTGACCGTGTCTCTCTAGTGTGAACCTGGGAACAATTC	142			

```
Qy 1950 CAAACTGTCCAGGGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 2009
|||||
Db 141 CAAACTGTCCAGGGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 82
|||||
Qy 2010 TCAGGGCCCATCCCTTCTGCGAGCTGTGACCAAAATTTAGTCCAGAAATAAATGAGA 2069
|||||
Db 81 TCAGGGCCCATCCCTTCTGCGAGCTGTGACCAAAATTTAGTCCAGAAATAAATGAGA 22
|||||
Qy 2070 AG 2071
||
Db 21 AG 20

RESULT 3
ABL63996/c
ID ABL63996 standard; DNA; 586 BP.
XX
AC ABL63996;
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:2333.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 26-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 29-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237423P.
```

```
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 2333; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ Sequence 586 BP; 130 A; 139 C; 179 G; 136 T; 2 other;
Query Match 23.3%; Score 482; DB 24; Length 586;
Best Local Similarity 100.0%; Pred. No. 5.5e-81;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1590 CAGGGTGATTCTGGGGGCGCTGGTGTGCAATGGCTCCCTGCAGGACATCGTGTCTCTGG 1649
|||||
DB 501 CAGGGTGATTCTGGGGGCGCTGGTGTGCAATGGCTCCCTGCAGGACATCGTGTCTCTGG 442
|||||
QY 1650 GGAGATTACCTTGTGCGCGCGCCCAACAGACACCGGGGTGTACACGAACCTCTGCAAGTTC 1709
|||||
DB 441 GGAGATTACCTTGTGCGCGCGCCCAACAGACACCGGGGTGTACACGAACCTCTGCAAGTTC 382
|||||
QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCGCAACTCCTGAGTATCCAGGACATCAGACACA 1769
|||||
DB 381 ACCAAGTGGATCCAGGAACCATCCAGGCGCAACTCCTGAGTATCCAGGACATCAGACACA 322
|||||
QY 1770 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCTCTTCAGACCCCTCATTCCTTC 1829
|||||
DB 321 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCTCTTCAGACCCCTCATTCCTTC 262
|||||
QY 1830 CCAGAGATGTTGAGATGTTTCATCTCTCAGCCCTGACCCCATGTCTCTGACACTCAGG 1889
|||||
DB 261 CCAGAGATGTTGAGATGTTTCATCTCTCAGCCCTGACCCCATGTCTCTGACACTCAGG 202
|||||
QY 1890 GTCTGCTTCCCACTGAGCTGACCGTGTCTCTCTAGTGTGAACCTCGGGAACAATTC 1949
|||||
DB 201 GTCTGCTTCCCACTGAGCTGACCGTGTCTCTCTAGTGTGAACCTCGGGAACAATTC 142
|||||
QY 1950 CAAAACTGTCCAGGCGGGGTTGGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 2009
|||||
DB 141 CAAAACTGTCCAGGCGGGGTTGGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 82
|||||
QY 2010 TCAGGGCCCATCCCTTCTCTGCGAGCTGTGACCCCAAAATTTAGTCCCAAGAAATAACTGAGA 2069
```

```
|||||
Db 81 TCAGGGCCCATCCCTCTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 22
QY 2070 AG 2071
Db 21 AG 20

RESULT 4
AAC79469/c
ID AAC79469 standard; cDNA; 735 BP.
XX AC
XX AAC79469;
XX 07-FEB-2001 (first entry)
XX cDNA sequence of human breast tumour clone B541S.
DE Human; breast tumour antigen; cytostatic; immunotherapy;
KW breast cancer; vaccine; ss.
XX Homo sapiens.
XX WO200061756-A2.
XX 19-OCT-2000.
XX 10-APR-2000; 2000WO-US09688.
XX 09-APR-1999; 99US-0288950.
XX 02-JUL-1999; 99US-0346327.
XX (CORI-) CORIXA CORP.
XX Reed SG, Xu J, Dillon DC;
PI WPI; 2000-638568/61.
XX A novel isolated polypeptide comprising an immunogenic portion of a
PT breast cancer protein useful in the detection and treatment of breast
PT cancer -
XX Claim 4; Page 88; 95pp; English.
XX The present sequence was isolated from a breast tumour cDNA library. It
XX is provided in a specification relating to compounds for immunotherapy
XX and diagnosis of breast cancer. Breast tumour antigens and the
XX polynucleotides that encode them may be used in the production of a
XX pharmaceutical composition to be used in the treatment of breast cancer.
XX Proliferated T cells and incubated antigen presenting cells are also
XX required. The polypeptides and polynucleotides may also be used to
XX produce a vaccine.
XX Sequence 735 BP; 161 A; 176 C; 226 G; 172 T; 0 other;
QY Query Match 23.3%; Score 482; DB 21; Length 735;
Best Local Similarity 100.0%; Pred. No. 5.7e-81;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1590 CAGGGTGATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 1649
Db 500 CAGGGTGATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 441
QY 1650 GGAGATTACCTTGTGCGGGGCCAACAGACGCGGGTGTCTACAGAACCTCTCAAGTTC 1709
Db 440 GGAGATTACCTTGTGCGGGGCCAACAGACGCGGGTGTCTACAGAACCTCTCAAGTTC 381
QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTAGTCAATCCAGGACTCAGCACA 1769
Db 380 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTAGTCAATCCAGGACTCAGCACA 321
QY 1770 CCGGCATCCCACTCTGTCGAGGAGACGCCCTGACACTCTCTTTCAGACCTCATTCCTTC 1829
|||||
```

```
Db 320 CCGGCATCCCACTCTGTCGAGGAGACGCCCTGACACTCTCTTTCAGACCTCATTCCTTC 261
QY 1830 CCAGAGATGTTGAGAAATGTTTCATCTCTCAGGCCCTGACCCCATGCTCTCTGACTCAGG 1889
Db 260 CCAGAGATGTTGAGAAATGTTTCATCTCTCAGGCCCTGACCCCATGCTCTCTGACTCAGG 201
QY 1890 GTCTGCTTCCCCACATTTGGGCTGACGCTGCTCTCTAGTTGAACCTCGGAAACAATTC 1949
Db 200 GTCTGCTTCCCCACATTTGGGCTGACGCTGCTCTCTAGTTGAACCTCGGAAACAATTC 141
QY 1950 CAAAACTGTCAGGGGGGGGTTGCGTCTCAATCTCCCTGGGACACTTTTCATCCTCAAGC 2009
Db 140 CAAAACTGTCAGGGGGGGGTTGCGTCTCAATCTCCCTGGGACACTTTTCATCCTCAAGC 81
QY 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 2069
Db 80 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 21
QY 2070 AG 2071
Db 20 AG 19

RESULT 5
ABK29013/c
ID ABK29013 standard; cDNA; 735 BP.
XX AC
XX ABK29013;
XX 23-APR-2002 (first entry)
XX Human breast tumour polypeptide cDNA clone #42.
DE Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
KW immunostimulant.
XX Homo sapiens.
XX WO200198339-A2.
XX 27-DEC-2001.
XX 12-JUN-2001; 2001WO-US19032.
XX 22-JUN-2000; 2000US-0602877.
XX 12-OCT-2000; 2000US-0687507.
XX 06-FEB-2001; 2001US-0778381.
XX (CORI-) CORIXA CORP.
XX Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;
PI WPI; 2002-147792/19.
XX Polynucleotides encoding breast tumour polypeptides, useful for treating
PT breast cancer or stimulating an immune response -
XX Claim 1; Page 140; 150pp; English.
XX The invention relates to polynucleotides encoding breast tumour
CC polypeptides. The sequences are useful for treating cancer, preferably
CC breast cancer, in a patient or for stimulating an immune response. The
CC polynucleotides and polypeptides are also useful in the diagnosis and
CC monitoring of breast cancer. A method for detecting the presence of a
CC cancer in a patient, comprises obtaining a biological sample from the
CC patient, contacting the biological sample with a binding agent that binds
CC to a breast tumour polypeptide, detecting in the sample an amount of
CC polypeptide that binds to the binding agent, and comparing the amount of
CC polypeptide to a predetermined cut-off value, therefore determining the
CC presence of a cancer in the patient. Sequences ABK28920-ABK29025
CC represent cDNA clones encoding human breast tumour polypeptides of the
CC invention.
XX
```



XX BS247; detection; diagnosis; breast cancer; atypical hyperplasia;  
 KW fibroadenoma; cystic breast disease; gene therapy; ss.  
 XX Homo sapiens.  
 OS WO9922027-A1.  
 XX 06-MAY-1999.  
 XX 28-OCT-1998; 98WO-US22906.  
 XX 28-OCT-1997; 97US-0968838.  
 XX 28-OCT-1997; 97US-0063431.  
 XX (ABBO ) ABBOTT LAB.  
 XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;  
 PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;  
 PI Russell JC, Stroupe SD, Yu H;  
 XX WPI; 1999-312977/26.  
 XX Breast tissue derived cDNA contig and consensus polypeptide sequence  
 PT Claim 1; Page 105; 112pp; English.  
 XX This sequence is a BS247 specific polynucleotide.  
 CC The invention relates to a method of detecting the presence of a target  
 CC BS247 polynucleotide, especially mRNA, in a test sample. BS247  
 CC polynucleotides are derived from breast tissue. The polynucleotides,  
 CC polypeptides or antibodies are useful for providing information leading  
 CC to the detection, diagnosis, staging, monitoring, prognosis, in vivo  
 CC imaging, prevention or treatment, determining predisposition to, diseases  
 CC and conditions of the breast, such as breast cancer, atypical  
 CC hyperplasia, fibroadenoma and cystic breast disease. Drug treatment or  
 CC gene therapy for breast cancer, can be based on these identified gene  
 CC sequences and the efficacy of any particular therapy can be monitored.  
 CC The BS247-derived reagents are advantageous for detection of breast  
 CC cancer due to their specificity. The reagents also provide an  
 CC alternative, non-surgical diagnostic method capable of detecting early  
 CC stage breast disease, such as cancer.  
 XX Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 other;  
 XX  
 Query Match 23.3%; Score 482; DB 20; Length 1302;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-81;  
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1590 CAGGGTGATTCCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 1649  
 DB 811 CAGGGTGATTCCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 870  
 QY 1650 GGAGATTACCTTGTGCCCCGGCCACACAGACCGGGGTGTACACGAACCTCTGCAAGTTC 1709  
 DB 871 GGAGATTACCTTGTGCCCCGGCCACACAGACCGGGGTGTACACGAACCTCTGCAAGTTC 930  
 QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGATCATCCAGGACTCAGCAC 1769  
 DB 931 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGATCATCCAGGACTCAGCAC 990  
 QY 1770 CCGGCATCCCACTGCTGTCAGGGACAGCCCTGACACTCTTTTCAGACCCCTCATTCCTTC 1829  
 DB 991 CCGGCATCCCACTGCTGTCAGGGACAGCCCTGACACTCTTTTCAGACCCCTCATTCCTTC 1050  
 QY 1830 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCCTCTGGACTCAGG 1889  
 DB 1051 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCCTCTGGACTCAGG 1110  
 QY 1890 GTCTGCTTCCCCACATTTGGGCTGACGGTGTCTCTAGTTGAACCCCTGGGAACAATTC 1949  
 DB 1111 GTCTGCTTCCCCACATTTGGGCTGACGGTGTCTCTAGTTGAACCCCTGGGAACAATTC 1170

QY 1950 CAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 2009  
 DB 1171 CAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1230  
 QY 2010 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCAAATTTAGTCCAGAAATAAACTCAGA 2069  
 DB 1231 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCAAATTTAGTCCAGAAATAAACTCAGA 1290  
 QY 2070 AG 2071  
 DB 1291 AG 1292  
 RESULT 8  
 AAX57990  
 ID AAX57990 standard; DNA; 1302 BP.  
 XX AC  
 XX AAX57990;  
 XX 19-JUL-1999 (first entry)  
 XX Human BS247 specific polynucleotide #8.  
 XX BS247; detection; diagnosis; breast cancer; atypical hyperplasia;  
 KW fibroadenoma; cystic breast disease; gene therapy; ss.  
 XX Homo sapiens.  
 XX WO9922027-A1.  
 XX 06-MAY-1999.  
 XX 28-OCT-1998; 98WO-US22906.  
 PR 28-OCT-1997; 97US-0968838.  
 PR 28-OCT-1997; 97US-0063431.  
 XX (ABBO ) ABBOTT LAB.  
 XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;  
 PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;  
 PI Russell JC, Stroupe SD, Yu H;  
 XX WPI; 1999-312977/26.  
 XX Breast tissue derived cDNA contig and consensus polypeptide sequence  
 PT Claim 1; Page 106; 112pp; English.  
 XX This sequence is a BS247 specific polynucleotide.  
 CC The invention relates to a method of detecting the presence of a target  
 CC BS247 polynucleotide, especially mRNA, in a test sample. BS247  
 CC polynucleotides are derived from breast tissue. The polynucleotides,  
 CC polypeptides or antibodies are useful for providing information leading  
 CC to the detection, diagnosis, staging, monitoring, prognosis, in vivo  
 CC imaging, prevention or treatment, determining predisposition to, diseases  
 CC and conditions of the breast, such as breast cancer, atypical  
 CC hyperplasia, fibroadenoma and cystic breast disease. Drug treatment or  
 CC gene therapy for breast cancer, can be based on these identified gene  
 CC sequences and the efficacy of any particular therapy can be monitored.  
 CC The BS247-derived reagents are advantageous for detection of breast  
 CC cancer due to their specificity. The reagents also provide an  
 CC alternative, non-surgical diagnostic method capable of detecting early  
 CC stage breast disease, such as cancer.  
 XX Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 other;  
 XX  
 Query Match 23.3%; Score 482; DB 20; Length 1302;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-81;  
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1590 CAGGGTGATTCCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 1649  
 DB 811 CAGGGTGATTCCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 870  
 QY 1650 GGAGATTACCTTGTGCCCCGGCCACACAGACCGGGGTGTACACGAACCTCTGCAAGTTC 1709  
 DB 871 GGAGATTACCTTGTGCCCCGGCCACACAGACCGGGGTGTACACGAACCTCTGCAAGTTC 930  
 QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGATCATCCAGGACTCAGCAC 1769  
 DB 931 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGATCATCCAGGACTCAGCAC 990  
 QY 1770 CCGGCATCCCACTGCTGTCAGGGACAGCCCTGACACTCTTTTCAGACCCCTCATTCCTTC 1829  
 DB 991 CCGGCATCCCACTGCTGTCAGGGACAGCCCTGACACTCTTTTCAGACCCCTCATTCCTTC 1050  
 QY 1830 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCCTCTGGACTCAGG 1889  
 DB 1051 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCCTCTGGACTCAGG 1110  
 QY 1890 GTCTGCTTCCCCACATTTGGGCTGACGGTGTCTCTAGTTGAACCCCTGGGAACAATTC 1949  
 DB 1111 GTCTGCTTCCCCACATTTGGGCTGACGGTGTCTCTAGTTGAACCCCTGGGAACAATTC 1170

Db 811 CAGGGTGATTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCTGG 870  
Qy 1650 GGAGATTACCTTTGTGCGGGGCCAACAGACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1709  
Db 871 GGAGATTACCTTTGTGCGGGGCCAACAGACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 930  
Qy 1710 ACCAAGTGGATCCAGGAACCAATCCAGGGCAACTCTGAGTCAATCCAGGACTCAGCACA 1769  
Db 931 ACCAAGTGGATCCAGGAACCAATCCAGGGCAACTCTGAGTCAATCCAGGACTCAGCACA 990  
Qy 1770 CCGGCTCCCACTGCTGCAGGACAGCGCTGCACACTCTCTTTCAGACCTCTATTCCTTC 1829  
Db 991 CCGGCTCCCACTGCTGCAGGACAGCGCTGCACACTCTCTTTCAGACCTCTATTCCTTC 1050  
Qy 1830 CCAGAGATGTTGAGATGTTCACTCTCTCCAGCGCCCTGACCCCATGTCTCTGAGCTCAGG 1889  
Db 1051 CCAGAGATGTTGAGATGTTCACTCTCTCCAGCGCCCTGACCCCATGTCTCTGAGCTCAGG 1110  
Qy 1890 GTCTGCTTCCCACTGCTGCAGGACAGCGCTGCACACTCTCTTTCAGACCTCTATTCCTTC 1949  
Db 1111 GTCTGCTTCCCACTGCTGCAGGACAGCGCTGCACACTCTCTTTCAGACCTCTATTCCTTC 1170  
Qy 1950 CAAAGTGTCCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCTCAAGC 2009  
Db 1171 CAAAGTGTCCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCTCAAGC 1230  
Qy 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAAAATTTAGTCCCAAGAAATAAAGTGA 2069  
Db 1231 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAAAATTTAGTCCCAAGAAATAAAGTGA 1290  
Qy 2070 AG 2071  
Db 1291 AG 1292

RESULT 9  
AAZ06259  
ID AAZ06259 standard; DNA; 1381 BP.  
XX AC AAZ06259;  
XX AC AAZ06259;  
XX DT 30-SEP-1999 (first entry)  
XX DE Human secreted protein gene No. 27.  
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
OS Homo sapiens.  
XX XX  
XX PN W09935158-A1.  
XX XX  
XX PD 15-JUL-1999.  
XX XX  
XX PF 06-JAN-1999; 99WO-US00108.  
XX PR 07-JAN-1998; 98US-0070704.  
XX PR 07-JAN-1998; 98US-0070657.  
XX PR 07-JAN-1998; 98US-0070658.  
XX PR 07-JAN-1998; 98US-0070692.  
XX XX  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;  
XX PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;  
XX XX  
XX DE WPI; 1999-444190/37.

DR P-PSDB; AAY38426.  
XX New isolated human genes and the secreted polypeptides they encode  
XX PS Claim 1; Page 173-174; 227pp; English.  
XX CC This sequence represents a nucleic acid molecule which encodes a  
CC secreted human protein. The gene number is given in the descriptor line.  
CC The gene can be used to generate fusion proteins by linking to the gene  
CC to a human immunoglobulin Fc portion (e.g. AAZ06210) for increasing the  
CC stability of the fused protein as compared to the human protein only.  
CC The invention relates to 36 novel genes and their fragments (nucleic  
CC acid sequences: AAZ06219-206263; amino acid sequences AAY38386-Y38498)  
CC which are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Also, pathological  
CC conditions can be diagnosed by determining the amount of the new  
CC polypeptides in a sample or by determining the presence of mutations in  
CC the new polynucleotides. Specific uses are described for each of the 36  
CC polynucleotides, based on which tissues they are most highly expressed in  
CC (see AAZ06219 for described uses).  
XX SQ Sequence 1381 BP; 329 A; 425 C; 353 G; 274 T; 0 other;  
Query Match 23.3%; Score 482; DB 20; Length 1381;  
Best Local Similarity 100.0%; Pred. No. 6.1e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1590 CAGGGTGATTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCTGG 1649  
Db 852 CAGGGTGATTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCTGG 911  
Qy 1650 GGAGATTACCTTTGTGCGGGGCCAACAGACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1709  
Db 912 GGAGATTACCTTTGTGCGGGGCCAACAGACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 971  
Qy 1710 ACCAAGTGGATCCAGGAACCAATCCAGGGCAACTCTCTGAGTCAATCCAGGACTCAGCACA 1769  
Db 972 ACCAAGTGGATCCAGGAACCAATCCAGGGCAACTCTCTGAGTCAATCCAGGACTCAGCACA 1031  
Qy 1770 CCGGCTCCCACTGCTGCAGGACAGCGCTGCACACTCTCTTTCAGACCTCTATTCCTTC 1829  
Db 1032 CCGGCTCCCACTGCTGCAGGACAGCGCTGCACACTCTCTTTCAGACCTCTATTCCTTC 1091  
Qy 1830 CCAGAGATGTTGAGATGTTCACTCTCTCCAGCGCCCTGACCCCATGTCTCTGAGCTCAGG 1889  
Db 1092 CCAGAGATGTTGAGATGTTCACTCTCTCCAGCGCCCTGACCCCATGTCTCTGAGCTCAGG 1151  
Qy 1890 GTCTGCTTCCCACTGCTGCAGGAGCGCTGCACACTCTCTCTAGTTGAACCTGGGAACAATTC 1949  
Db 1152 GTCTGCTTCCCACTGCTGCAGGAGCGCTGCACACTCTCTCTAGTTGAACCTGGGAACAATTC 1211  
Qy 1950 CAAAGTGTCCAGGCGGGGTTGCGTCTCAATCTCCCTGGGGAGCTTTTCATCTCTCAAGC 2009  
Db 1212 CAAAGTGTCCAGGCGGGGTTGCGTCTCAATCTCCCTGGGGAGCTTTTCATCTCTCAAGC 1271  
Qy 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAAAATTTAGTCCCAAGAAATAAAGTGA 2069  
Db 1272 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAAAATTTAGTCCCAAGAAATAAAGTGA 1331  
Qy 2070 AG 2071  
Db 1332 AG 1333  
RESULT 10  
AAZ06260  
ID AAZ06260 standard; DNA; 1439 BP.  
XX AC AAZ06260;  
XX AC AAZ06260;  
XX DT 30-SEP-1999 (first entry)  
XX XX  
XX DE Human secreted protein gene No. 27.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
 KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

OS Homo sapiens.

XX WO9935158-A1.

PD 15-JUL-1999.

PF 06-JAN-1999; 99WO-US00108.

XX 07-JAN-1998; 98US-0070704.

PR 07-JAN-1998; 98US-0070657.

PR 07-JAN-1998; 98US-0070658.

PR 07-JAN-1998; 98US-0070692.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;  
 PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;

XX WPI; 1999-444190/37.

DR P-PSDB; AAY38427.

XX New isolated human genes and the secreted polypeptides they encode

PS Claim 1; Page 174; 227pp; English.

XX This sequence represents a nucleic acid molecule which encodes a  
 CC secreted human protein. The gene number is given in the descriptor line.  
 CC The gene can be used to generate fusion proteins by linking to the gene  
 CC to a human immunoglobulin Fc portion (e.g. AAZ06210) for increasing the  
 CC stability of the fused protein as compared to the human protein only.  
 CC The invention relates to 36 novel genes and their fragments (nucleic  
 CC acid sequences: AAZ06219-206263; amino acid sequences AAY38386-Y38498)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 36  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAZ06219 for described uses).

XX Sequence 1439 BP; 325 A; 446 C; 355 G; 313 T; 0 other;

Query Match 23.3%; Score 482; DB 20; Length 1439;

Best Local Similarity 100.0%; Pred. No. 6.2e-81;

Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGTGATTCGGGGGCGCTGTGCTGCAATGCTCCCTGCAGGACTCGTGCTCGG 1649

Db 910 CAGGTGATTCGGGGGCGCTGTGCTGCAATGCTCCCTGCAGGACTCGTGCTCGG 969

QY 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACGGGTGTCTACACGAACCTCTGCAAGTTC 1709

Db 970 GGAGATTACCTTGTGCGCGGCCCAACAGACGGGTGTCTACACGAACCTCTGCAAGTTC 1029

QY 1710 ACCAAGTGGATCCAGGAACCAATCCAGGCCAACTCCTGAGTATCCAGGACTAGCACA 1769

Db 1030 ACCAAGTGGATCCAGGAACCAATCCAGGCCAACTCCTGAGTATCCAGGACTAGCACA 1089

QY 1770 CGGGATCCCACTGCTGCGGGGACAGCCCTGACACTCCCTTCAGACCTCATTCCTTC 1829

Db 1090 CGGGATCCCACTGCTGCGGGGACAGCCCTGACACTCCCTTCAGACCTCATTCCTTC 1149

QY 1830 CCAGAGATGTTGAGAATGTTATCTCTCCAGCCCTGACCCCATCTCTCTGGACTCAGG 1889

Db 1150 CCAGAGATGTTGAGAATGTTATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1209

QY 1890 GTCTGCTTCCCCACACATTTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGACAAATTC 1949

Db 1210 GTCTGCTTCCCCACACATTTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGACAAATTC 1269

QY 1950 CAAAACTGTCCAGGCGGGGGTGGGTCTCAATCTCCCTGGGGCACTTTTCATCCTCAAGC 2009

Db 1270 CAAAACTGTCCAGGCGGGGGTGGGTCTCAATCTCCCTGGGGCACTTTTCATCCTCAAGC 1329

QY 2010 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCGAGAATAAATCTGAGA 2069

Db 1330 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCGAGAATAAATCTGAGA 1389

QY 2070 AG 2071

Db 1390 AG 1391

RESULT 11

AA60578

ID AAX60578 standard; DNA; 1499 BP.

XX AAX60578;

AC AAX60578;

XX 02-AUG-1999 (first entry)

DT Human keratinocyte derived protease (KDP) encoding DNA.

XX Keratinocyte derived protease; KDP; protease; keratinocyte; human;

KW skin care product; skin flaking; dandruff; laundry detergent composition;

KW cleaning composition; dishwashing product; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT 5'UTR 1..290

FT CDS /tag= a

FT 291..1172

FT /product= "keratinocyte derived protease"

FT sig\_peptide 291..488

FT /tag= c

FT mat\_peptide 489..1169

FT /tag= d

FT 3'UTR 1173..1499

FT /tag= e

XX WO9918219-A1.

XX 15-APR-1999.

XX 03-OCT-1997; 97WO-US17864.

XX 03-OCT-1997; 97WO-US17864.

XX (PROC ) PROCTER & GAMBLE CO.

XX Kitado H, Yoshikawa A, Zaiki T;

XX WPI; 1999-287737/24.

XX P-PSDB; AAY16777.

XX Novel human protease useful for treating or preventing skin flaking

XX Claim 3; Page 32-34; 43pp; English.

XX This DNA encodes a human keratinocyte derived protease (KDP) polypeptide.

XX The protease is substantially similar to a sequence encoded by the insert

XX in plasmid FERM BP-6129. The KDP polypeptide, from human keratinocytes,

XX is specifically used in skin care products, particularly to treat or

XX prevent skin flaking, e.g. dandruff, and in laundry detergent

CC compositions. More generally the KDP polypeptide can be used in any  
CC cleaning composition, e.g. hard surface or dishwashing products.  
XX  
SQ Sequence 1499 BP; 310 A; 469 C; 398 G; 322 T; 0 other;

Query Match 23.3%; Score 482; DB 20; Length 1499;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1590 CAGGGTATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTCTCTGG 1649  
Db 1014 CAGGGTATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTCTCTGG 1073

Qy 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1709  
Db 1074 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1133

Qy 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCCACTCTGAGTCAATCCAGGACTCAGCACA 1769  
Db 1134 ACCAAGTGGATCCAGGAACCATCCAGGCCCACTCTGAGTCAATCCAGGACTCAGCACA 1193

Qy 1770 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTCAGACCTCATTCCTTC 1829  
Db 1194 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTCAGACCTCATTCCTTC 1253

Qy 1830 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1889  
Db 1254 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1313

Qy 1890 GTCTGCTTCCCACTGCTGCAGGACAGCCCTGACACTCCTTTCAGACCTCATTCCTTC 1949  
Db 1314 GTCTGCTTCCCACTGCTGCAGGACAGCCCTGACACTCCTTTCAGACCTCATTCCTTC 1373

Qy 1950 CAAAACGTCCAGGCGGGGTGCGTCTCAATCTCCCTGGGACATTCATCTCTCAAGC 2009  
Db 1374 CAAAACGTCCAGGCGGGGTGCGTCTCAATCTCCCTGGGACATTCATCTCTCAAGC 1433

Qy 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCAGGA 2069  
Db 1434 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCAGGA 1493

Qy 2070 AG 2071  
Db 1494 AG 1495

RESULT 12  
AAZ06245  
XX AAZ06245 standard; DNA; 1516 BP.  
XX  
XX  
XX  
XX 30-SEP-1999 (first entry)  
XX  
DE Human secreted protein gene No. 27.  
XX  
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizoprenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX  
OS Homo sapiens.  
XX  
XX WO9935158-A1.  
XX  
XX 15-JUL-1999.  
XX  
XX 06-JAN-1999; 99WO-US00108.  
XX  
XX

PR 07-JAN-1998; 98US-0070704.  
PR 07-JAN-1998; 98US-0070657.  
PR 07-JAN-1998; 98US-0070658.  
PR 07-JAN-1998; 98US-0070692.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Brewer LA, Duan RD, Ebner R, Lafleur DW, NI J;  
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;  
XX  
DR WPI; 1999-444190/37.  
DR P-PSDB; AAY38412.  
XX  
PT New isolated human genes and the secreted polypeptides they encode  
XX  
XX Claim 1; Page 163-164; 227pp; English.  
XX  
CC This sequence represents a nucleic acid molecule which encodes a  
CC secreted human protein. The gene number is given in the descriptor line.  
CC The gene can be used to generate fusion proteins by linking to the gene  
CC to a human immunoglobulin Fc portion (e.g. AA206210) for increasing the  
CC stability of the fused protein as compared to the human protein only.  
CC The invention relates to 36 novel genes and their fragments (nucleic  
CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)  
CC which are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Also, pathological  
CC polypeptides in a sample or by determining the amount of the new  
CC the new polynucleotides. Specific uses are described for each of the 36  
CC polynucleotides, based on which tissues they are most highly expressed in  
CC (see AA206219 for described uses).  
XX  
SQ Sequence 1516 BP; 338 A; 472 C; 381 G; 324 T; 1 other;

Query Match 23.3%; Score 482; DB 20; Length 1516;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1590 CAGGGTATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTCTCTCTGG 1649  
Db 977 CAGGGTATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTCTCTCTGG 1036

Qy 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1709  
Db 1037 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1096

Qy 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCCACTCTGAGTCAATCCAGGACTCAGCACA 1769  
Db 1097 ACCAAGTGGATCCAGGAACCATCCAGGCCCACTCTGAGTCAATCCAGGACTCAGCACA 1156

Qy 1770 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCTCATTCCTTC 1829  
Db 1157 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCTCATTCCTTC 1216

Qy 1830 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1889  
Db 1217 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1276

Qy 1890 GTCTGCTTCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCTCATTCCTTC 1949  
Db 1277 GTCTGCTTCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCTCATTCCTTC 1336

Qy 1950 CAAAACGTCCAGGCGGGGTGCGTCTCAATCTCCCTGGGACATTCATCTCTCAAGC 2009  
Db 1337 CAAAACGTCCAGGCGGGGTGCGTCTCAATCTCCCTGGGACATTCATCTCTCAAGC 1396

Qy 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCAGGA 2069  
Db 1397 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCAGGA 1456

Qy 2070 AG 2071  
Db 1457 AG 1458



## RESULT 13

AAZ65070

ID AAZ65070 standard; cDNA; 1570 BP.

XX

AC AAZ65070;

XX

DT 05-APR-2000 (first entry)

XX

DE Membrane-bound protein PRO1132 encoding cDNA.

XX

KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;

KW pharmaceutical; receptor immunoadhesin; gene mapping; ss.

XX

OS Homo sapiens.

XX

PN WO9963088-A2.

XX

PD 09-DEC-1999.

XX

PF 02-JUN-1999; 99WO-US12252.

XX

PR 02-JUN-1998; 98US-0087607.

PR

PR 02-JUN-1998; 98US-0087609.

PR

PR 02-JUN-1998; 98US-0087759.

PR

PR 03-JUN-1998; 98US-0087827.

PR

PR 04-JUN-1998; 98US-0088021.

PR

PR 04-JUN-1998; 98US-0088025.

PR

PR 04-JUN-1998; 98US-0088028.

PR

PR 04-JUN-1998; 98US-0088029.

PR

PR 04-JUN-1998; 98US-0088030.

PR

PR 04-JUN-1998; 98US-0088033.

PR

PR 04-JUN-1998; 98US-0088326.

PR

PR 05-JUN-1998; 98US-0088167.

PR

PR 05-JUN-1998; 98US-0088202.

PR

PR 05-JUN-1998; 98US-0088212.

PR

PR 05-JUN-1998; 98US-0088217.

PR

PR 05-JUN-1998; 98US-0088555.

PR

PR 10-JUN-1998; 98US-0088722.

PR

PR 10-JUN-1998; 98US-0088730.

PR

PR 10-JUN-1998; 98US-0088734.

PR

PR 10-JUN-1998; 98US-0088738.

PR

PR 10-JUN-1998; 98US-0088740.

PR

PR 10-JUN-1998; 98US-0088741.

PR

PR 10-JUN-1998; 98US-0088742.

PR

PR 10-JUN-1998; 98US-0088810.

PR

PR 10-JUN-1998; 98US-0088811.

PR

PR 10-JUN-1998; 98US-0088824.

PR

PR 10-JUN-1998; 98US-0088825.

PR

PR 10-JUN-1998; 98US-0088826.

PR

PR 11-JUN-1998; 98US-0088858.

PR

PR 11-JUN-1998; 98US-0088861.

PR

PR 11-JUN-1998; 98US-0088863.

PR

PR 11-JUN-1998; 98US-0088876.

PR

PR 12-JUN-1998; 98US-0089090.

PR

PR 12-JUN-1998; 98US-0089105.

PR

PR 16-JUN-1998; 98US-0089440.

PR

PR 16-JUN-1998; 98US-0089512.

PR

PR 16-JUN-1998; 98US-0089514.

PR

PR 17-JUN-1998; 98US-0089532.

PR

PR 17-JUN-1998; 98US-0089538.

PR

PR 17-JUN-1998; 98US-0089598.

PR

PR 17-JUN-1998; 98US-0089599.

PR

PR 17-JUN-1998; 98US-0089600.

PR

PR 17-JUN-1998; 98US-0089653.

PR

PR 18-JUN-1998; 98US-0089801.

PR

PR 18-JUN-1998; 98US-0089907.

PR

PR 18-JUN-1998; 98US-0089908.

PR

PR 18-JUN-1998; 98US-0089908.

PR

PR 19-JUN-1998; 98US-0089947.

PR

PR 19-JUN-1998; 98US-0089948.

PR

PR 19-JUN-1998; 98US-0089952.

PR

PR 22-JUN-1998; 98US-0090246.

PR

PR 22-JUN-1998; 98US-0090252.  
PR 22-JUN-1998; 98US-0090254.  
PR 23-JUN-1998; 98US-0090349.  
PR 23-JUN-1998; 98US-0090355.  
PR 24-JUN-1998; 98US-0090429.  
PR 24-JUN-1998; 98US-0090431.  
PR 24-JUN-1998; 98US-0090435.  
PR 24-JUN-1998; 98US-0090444.  
PR 24-JUN-1998; 98US-0090445.  
PR 24-JUN-1998; 98US-0090461.  
PR 24-JUN-1998; 98US-0090472.  
PR 24-JUN-1998; 98US-0090535.  
PR 24-JUN-1998; 98US-0090538.  
PR 24-JUN-1998; 98US-0090540.  
PR 24-JUN-1998; 98US-0090557.  
PR 25-JUN-1998; 98US-0090676.  
PR 25-JUN-1998; 98US-0090678.  
PR 25-JUN-1998; 98US-0090688.  
PR 25-JUN-1998; 98US-0090690.  
PR 25-JUN-1998; 98US-0090691.  
PR 25-JUN-1998; 98US-0090694.  
PR 25-JUN-1998; 98US-0090695.  
PR 25-JUN-1998; 98US-0090696.  
PR 26-JUN-1998; 98US-0090862.  
PR 26-JUN-1998; 98US-0090863.  
PR 01-JUL-1998; 98US-0091358.  
PR 01-JUL-1998; 98US-0091360.  
PR 01-JUL-1998; 98US-0091344.  
PR 02-JUL-1998; 98US-0091478.  
PR 02-JUL-1998; 98US-0091486.  
PR 02-JUL-1998; 98US-0091519.  
PR 02-JUL-1998; 98US-0091626.  
PR 02-JUL-1998; 98US-0091628.  
PR 02-JUL-1998; 98US-0091633.  
PR 02-JUL-1998; 98US-0091646.  
PR 02-JUL-1998; 98US-0091673.  
PR 07-JUL-1998; 98US-0091978.  
PR 07-JUL-1998; 98US-0091982.  
PR 09-JUL-1998; 98US-0092182.  
PR 10-JUL-1998; 98US-0092472.  
PR 20-JUL-1998; 98US-0093339.  
PR 30-JUL-1998; 98US-0094651.  
PR 04-AUG-1998; 98US-0095282.  
PR 04-AUG-1998; 98US-0095285.  
PR 04-AUG-1998; 98US-0095301.  
PR 04-AUG-1998; 98US-0095302.  
PR 04-AUG-1998; 98US-0095318.  
PR 04-AUG-1998; 98US-0095321.  
PR 04-AUG-1998; 98US-0095325.  
PR 10-AUG-1998; 98US-0095916.  
PR 10-AUG-1998; 98US-0095929.  
PR 10-AUG-1998; 98US-0096012.  
PR 11-AUG-1998; 98US-0096143.  
PR 11-AUG-1998; 98US-0096146.  
PR 12-AUG-1998; 98US-0096329.  
PR 17-AUG-1998; 98US-0096757.  
PR 17-AUG-1998; 98US-0096766.  
PR 17-AUG-1998; 98US-0096768.  
PR 17-AUG-1998; 98US-0096773.  
PR 17-AUG-1998; 98US-0096791.  
PR 17-AUG-1998; 98US-0096867.  
PR 17-AUG-1998; 98US-0096891.  
PR 17-AUG-1998; 98US-0096894.  
PR 17-AUG-1998; 98US-0096895.  
PR 18-AUG-1998; 98US-0096897.  
PR 18-AUG-1998; 98US-0096949.  
PR 18-AUG-1998; 98US-0096950.  
PR 18-AUG-1998; 98US-0096959.  
PR 18-AUG-1998; 98US-0096960.  
PR 18-AUG-1998; 98US-0097022.  
PR 19-AUG-1998; 98US-0097141.  
PR 20-AUG-1998; 98US-0097218.  
PR 24-AUG-1998; 98US-0097661.

PR 26-AUG-1998; 98US-0097951.  
PR 26-AUG-1998; 98US-0097952.  
PR 26-AUG-1998; 98US-0097954.  
PR 26-AUG-1998; 98US-0097955.  
PR 26-AUG-1998; 98US-0097971.  
PR 26-AUG-1998; 98US-0097974.  
PR 26-AUG-1998; 98US-0097978.  
PR 26-AUG-1998; 98US-0097979.  
PR 26-AUG-1998; 98US-0097986.  
PR 26-AUG-1998; 98US-0098014.  
PR 31-AUG-1998; 98US-0098525.  
PR 16-SEP-1998; 98US-0100634.  
PR 12-JAN-1999; 99US-0115365.  
XX XX  
PA (GETH ) GENENTECH INC.  
XX XX  
PI Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;  
PI Wood WI, Yuan J;  
XX XX  
DR WPI; 2000-072883/06.  
DR P-PSDB; AAY66726.  
XX XX  
PT Membrane-bound proteins and related nucleotide sequences -  
XX XX  
PS Claim 2; Fig 225; 822pp; English.  
XX XX  
CC The invention provides membrane-bound PRO polypeptides and  
CC polynucleotides encoding them. The PRO sequences of the invention were  
CC identified based on extracellular domain homology screening. The PRO  
CC sequences have homology with proteins including LDL receptors, TIE  
CC ligands and various enzymes. The membrane-bound proteins and receptor  
CC molecules are useful as pharmaceutical and diagnostic agents. Receptor  
CC immunoadhesins, for instance, can be used as therapeutic agents to block  
CC receptor-ligand interactions. The membrane-bound proteins can also be  
CC employed for screening of potential peptide or small molecule inhibitors  
CC of the relevant receptor/ligand interaction. The PRO encoding sequences  
CC are useful as hybridization probes, in chromosome and gene mapping and in  
CC the generation of antisense RNA and DNA. PRO nucleic acid sequences  
CC will also be useful for the preparation of PRO polypeptides, especially  
CC by recombinant techniques.  
XX XX  
SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

Query Match 23.3%; Score 482; DB 21; Length 1570;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGCTGATCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1649  
DB 1077 CAGGCTGATCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1136  
QY 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1709  
DB 1137 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1196  
QY 1710 ACCAAGTGGATCCAGGAACCAATCCAGGCCAACTCTGAGTCATCCAGGACTCAGCACA 1769  
DB 1197 ACCAAGTGGATCCAGGAACCAATCCAGGCCAACTCTGAGTCATCCAGGACTCAGCACA 1256  
QY 1770 CCGGCATCCCACTGTGTGAGGACAGCCCTGTACACTCTCTTCAGACCTCATTCCTTC 1829  
DB 1257 CCGGCATCCCACTGTGTGAGGACAGCCCTGTACACTCTCTTCAGACCTCATTCCTTC 1316  
QY 1830 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCGGACTCAGG 1889  
DB 1317 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCGGACTCAGG 1376  
QY 1890 GTCTGCTTCCCACTATGGGCTGACCGTGTCTCTAGTTGAACCTCGGGAACAATTC 1949  
DB 1377 GTCTGCTTCCCACTATGGGCTGACCGTGTCTCTAGTTGAACCTCGGGAACAATTC 1436  
QY 1950 CAAACTGTCCAGGCGGGGTTGGTCTCAATCTCCCTGGGCACTTTCATCTCCTCAGC 2009

Db 1437 CAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1496  
QY 2010 TCAGGCCCCATCCCTTCTCTGACAGCTCTGACCCAAATTTAGTCCCAAGATAAATGAGA 2069  
Db 1497 TCAGGCCCCATCCCTTCTCTGACAGCTCTGACCCAAATTTAGTCCCAAGATAAATGAGA. 1556  
QY 2070 AG 2071  
Db 1557 AG 1558  
RESULT 14  
AAS21471  
ID AAS21471 standard; cDNA; 1570 BP.  
XX XX  
AC AAS21471;  
XX XX  
DT 24-OCT-2001 (first entry)  
XX XX  
DE Human cDNA sequence encoding for PRO1132 polypeptide.  
XX XX  
KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.  
XX OS  
XX Homo sapiens.  
XX PN WO200140466-A2.  
XX PD 07-JUN-2001.  
XX PF 01-DEC-2000; 2000WO-US32678.  
XX PR 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28584.  
PR 02-DEC-1999; 99WO-US28585.  
PR 09-DEC-1999; 99US-0170262.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 20-DEC-1999; 99WO-US30999.  
PR 30-DEC-1999; 99WO-US31243.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 10-NOV-2000; 2000WO-US30873.  
XX XX  
PA (GETH ) GENENTECH INC.  
XX XX  
PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX WPI; 2001-408281/43.  
DR P-PSDB; AAU12399.  
XX XX  
PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
other PRO polypeptides, link bioactive molecules to cells expressing

PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
PT lung, breast, prostate, cervical  
PS Claim 3; Fig 455; 813pp; English.  
XX  
XX  
XX AAS21244-AAS21518 encode for novel human secretory and transmembrane  
CC PRO polypeptides. The PRO polypeptides are useful to detect other  
CC PRO polypeptides, to link bioactive molecules to cells expressing  
CC PRO polypeptides, to modulate biological activities of cells expressing  
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
CC polypeptide expression in a cell sample to that in a control sample.  
CC Some of the 275 sequences are also useful to stimulate the release of  
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
CC proliferation or differentiation of chondrocytes, the proliferation or  
CC gene expression in pericyte cells, the release of proteoglycans from  
CC cartilage, the proliferation of inner ear utricular supporting cells or  
CC of T-lymphocytes, the release of a cytokine from peripheral blood  
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
CC molecules involved in binding interactions. The polynucleotides encoding  
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
CC transgenic or knock out animals and can be used in gene therapy.  
XX  
XX Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

Query Match 23.3%; Score 482; DB 22; Length 1570;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1590 CAGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCTCAGGAGCTCTGTCCTGG 1649  
Db 1077 CAGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCTCAGGAGCTCTGTCCTGG 1136  
QY 1650 GGAGATTACCTTGTGCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1709  
Db 1137 GGAGATTACCTTGTGCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1196  
QY 1710 ACCAAGTGATCCAGGAACCATCCAGGCCAATCCTGAGTCTATCCAGGACTCAGCACA 1769  
Db 1197 ACCAAGTGATCCAGGAACCATCCAGGCCAATCCTGAGTCTATCCAGGACTCAGCACA 1256  
QY 1770 CCGGATCCCACTGCTCAGGGACAGCCCTGACACTCCTTTGACACCTCATCTCCTTC 1829  
Db 1257 CCGGATCCCACTGCTCAGGGACAGCCCTGACACTCCTTTGACACCTCATCTCCTTC 1316  
QY 1830 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1889  
Db 1317 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1376  
QY 1890 GTCTGCTCCCAATGAGGCTGACCGTGTCTCTAGTCTAGTGAACCTGGGAACATTC 1949  
Db 1377 GTCTGCTCCCAATGAGGCTGACCGTGTCTCTAGTCTAGTGAACCTGGGAACATTC 1436  
QY 1950 CAAAAGTCTCAGGGGGGGTGGCTCTCAATCTCCCTGGGCACTTTCATCTCCTCAAGC 2009  
Db 1437 CAAAAGTCTCAGGGGGGGTGGCTCTCAATCTCCCTGGGCACTTTCATCTCCTCAAGC 1496  
QY 2010 TCAGGGCCCATCCCTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 2069  
Db 1497 TCAGGGCCCATCCCTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 1556  
QY 2070 AG 2071  
Db 1557 AG 1558

RESULT 15  
AAF44216  
ID AAF44216 standard; cDNA; 1570 BP.  
XX

AC AAF44216;  
XX  
XX 02-APR-2001 (first entry)  
XX  
DE Human PRO1132 (UN0570) nucleotide sequence SEQ ID NO:308.  
XX  
XX Human; secreted and transmembrane protein; PRO; cytotstatic;  
KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;  
KW diagnostic assay; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200073454-A1.  
XX  
PD 07-DEC-2000.  
XX  
PF 30-MAR-2000; 2000WO-US08439.  
XX  
PR 02-JUN-1999; 99WO-US12252.  
PR 23-JUN-1999; 99US-0141037.  
PR 07-JUL-1999; 99US-0143048.  
PR 20-JUL-1999; 99US-0144758.  
PR 26-JUL-1999; 99US-0145698.  
PR 28-JUL-1999; 99US-0146222.  
PR 17-AUG-1999; 99US-0149396.  
PR 15-SEP-1999; 99WO-US21090.  
PR 08-OCT-1999; 99US-0158663.  
PR 30-NOV-1999; 99WO-US28313.  
PR 01-DEC-1999; 99WO-US28301.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 02-MAR-2000; 2000WO-US05004.  
PR 15-MAR-2000; 2000WO-US05841.  
PR 20-MAR-2000; 2000WO-US06884.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
PI Ferrera N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;  
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;  
PI Zhang Z;  
XX  
XX WPI: 2001-032160/04.  
DR P-PSDB; AAB65249.  
XX  
PT PRO polynucleotides used to produce polypeptides used to target  
PT bioactive molecules such as toxins, radiolabels or antibodies, to  
PT specific cells, to cause targeted cell death -  
XX  
XX Claim 2; Fig 225; 935pp; English.  
XX  
XX The present invention describes human secreted and transmembrane PRO  
CC proteins. The PRO proteins have cytotstatic activity. The PRO proteins  
CC can be used for targeted delivery of bioactive molecules, such as  
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide  
CC sequences, and their fragments, can be used as hybridisation probes, in  
CC chromosomal and gene mapping, and in the generation of anti-sense RNA  
CC and DNA. They may also be used to produce transgenic animals which are  
CC used to develop and screen therapeutically useful reagents. The PRO  
CC nucleotide and protein sequence can be used for tissue typing and in  
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.  
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used  
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and  
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein

CC sequences given in the exemplification of the present invention.

XX  
SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;  
Query Match 23.3%; Score 482; DB 22; Length 1570;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1590 CAGGGTGATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1649  
Db 1077 CAGGGTGATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1136  
Qy 1650 GGAGATTACCTTGTGCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1709  
Db 1137 GGAGATTACCTTGTGCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1196  
Qy 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 1769  
Db 1197 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 1256  
Qy 1770 CCGGCATCCCACTCTGTCAGGGACAGCCCTGACACTCCTTTACAGACCTCATTCCTTC 1829  
Db 1257 CCGGCATCCCACTCTGTCAGGGACAGCCCTGACACTCCTTTACAGACCTCATTCCTTC 1316  
Qy 1830 CCAGAGATGTGAGATGTTCATCTCTCCAGCCCTTGACCCCATGTCTCTGGACTCAGG 1889  
Db 1317 CCAGAGATGTGAGATGTTCATCTCTCCAGCCCTTGACCCCATGTCTCTGGACTCAGG 1376  
Qy 1890 GTCTGCTTCCCCACATTTGGCTGACCGTGTCTCTAGTTGAACCTGGGAACAATTC 1949  
Db 1377 GTCTGCTTCCCCACATTTGGCTGACCGTGTCTCTAGTTGAACCTGGGAACAATTC 1436  
Qy 1950 CAAACTGTCCAGGGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 2009  
Db 1437 CAAACTGTCCAGGGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1496  
Qy 2010 TCAGGGCCCATCCCTTCTCTGACAGCTCTGACCCAAATTTAGTCCAGAAATAAACTGAGA 2069  
Db 1497 TCAGGGCCCATCCCTTCTCTGACAGCTCTGACCCAAATTTAGTCCAGAAATAAACTGAGA 1556  
Qy 2070 AG 2071  
Db 1557 AG 1558

Search completed: August 5, 2003, 05:55:17  
Job time : 513.264 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 5, 2003, 05:12:55 ; Search time 134.302 Seconds  
(without alignments)  
6806.329 Million cell updates/sec

Title: US-09-936-271b-13\_COPY\_9500\_11570  
Perfect score: 2071  
Sequence: 1 aaacagacacaaacttctt.....cccagaaataaactgagaag 2071

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	482	23.3	735	US-09-602-877A-94	Sequence 94, Appl
C 2	482	23.3	1570	US-09-996-243-308	Sequence 308, App
C 3	476	23.0	1476	US-08-824-874-2	Sequence 2, Appli
C 4	476	23.0	1476	US-09-210-084-2	Sequence 2, Appli
C 5	476	23.0	1476	US-09-764-762-2	Sequence 2, Appli
C 6	460	22.2	1504	US-09-280-116-1	Sequence 1, Appli
C 7	222.6	10.7	176373	US-09-128-153-17	Sequence 17, Appl
C 8	219.4	10.6	70000	US-09-851-896-3	Sequence 3, Appli
C 9	219.2	10.6	152331	US-09-128-155-16	Sequence 16, Appl
C 10	218.6	10.6	25464	US-09-326-480A-4	Sequence 4, Appli
C 11	216.4	10.4	461	US-09-404-879A-1	Sequence 1, Appli
C 12	216.4	10.4	461	US-09-404-879A-3	Sequence 3, Appli
C 13	216.4	10.4	461	US-09-338-933-1	Sequence 1, Appli
C 14	216.4	10.4	461	US-09-338-933-3	Sequence 3, Appli
C 15	216.4	10.4	461	US-09-215-681-1	Sequence 1, Appli
C 16	216.4	10.4	461	US-09-215-681-3	Sequence 3, Appli
C 17	215.6	10.4	3460	US-09-904-615-44	Sequence 44, Appl
C 18	214.6	10.4	1001	US-09-671-317-170	Sequence 170, App
C 19	214.6	10.4	1001	US-09-671-317-450	Sequence 450, App
C 20	214.6	10.4	41684	US-09-536-059-1	Sequence 1, Appli
C 21	212.6	10.3	4421	US-08-257-963B-9	Sequence 9, Appli
C 22	212.6	10.3	4421	US-08-367-841A-9	Sequence 9, Appli
C 23	212.6	10.3	4421	US-08-520-373D-6	Sequence 6, Appli
C 24	212.6	10.3	4421	PCT-US95-07201-9	Sequence 9, Appli
C 25	212.6	10.3	148567	US-09-801-876B-3	Sequence 3, Appli
C 26	212.2	10.2	99500	US-09-798-096-10	Sequence 10, Appl
C 27	210.4	10.2	38553	US-09-922-445-1	Sequence 1, Appli

C 28	209.8	10.1	55827	4	US-09-813-133A-3	Sequence 3, Appli
C 29	209.2	10.1	66804	4	US-09-740-041-3	Sequence 3, Appli
C 30	209	10.1	15328	2	US-08-888-497-33	Sequence 33, Appl
C 31	209	10.1	15328	4	US-09-362-230-33	Sequence 33, Appl
C 32	209	10.1	15328	5	PCT-US94-07926-33	Sequence 33, Appl
C 33	209	10.1	36651	4	US-09-738-894A-3	Sequence 3, Appli
C 34	209	10.1	36651	4	US-09-964-469-3	Sequence 3, Appli
C 35	209	10.1	49312	4	US-09-671-317-485	Sequence 485, App
C 36	207.4	10.0	6088	4	US-09-620-312D-190	Sequence 190, App
C 37	207.2	10.0	319608	4	US-09-539-333D-1	Sequence 1, Appli
C 38	207.2	10.0	319608	4	US-09-679-409-1	Sequence 1, Appli
C 39	207	10.0	9365	4	US-09-608-285A-8	Sequence 8, Appli
C 40	207	10.0	9365	4	US-09-350-836B-8	Sequence 8, Appli
C 41	207	10.0	9365	4	US-09-370-265-8	Sequence 8, Appli
C 42	207	10.0	9365	4	US-09-557-800C-8	Sequence 8, Appli
C 43	207	10.0	14581	4	US-08-520-373D-4	Sequence 4, Appli
C 44	207	10.0	14747	4	US-09-608-285A-42	Sequence 42, Appl
C 45	207	10.0	14747	4	US-09-557-800C-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1

US-09-602-877A-94/c  
; Sequence 94, Application US/09602877A  
; Patent No. 6432707  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.446C5  
; CURRENT APPLICATION NUMBER: US/09/602,877A  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 94  
; LENGTH: 735  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-602-877A-94

Query Match	23.3%	Score 482;	DB 4;	Length 735;
Best Local Similarity	100.0%	Pred. No. 6.7e-92;		
Matches 482;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1590	CAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCGCAGGACTCGTGTCTGG 1649		
DB	500	CAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCGCAGGACTCGTGTCTGG 441		
QY	1650	GGAGATTACCTTTGTGCCCGGCCCAACAGACGGGGTGTCTACACGAACCTCTGCAAGTTC 1709		
DB	440	GGAGATTACCTTTGTGCCCGGCCCAACAGACGGGGTGTCTACACGAACCTCTGCAAGTTC 381		
QY	1710	ACCAAGTGATCCAGGAACCATCCAGGCCACTCTGTAGTCTATCCAGGACTCAGACA 1769		
DB	380	ACCAAGTGATCCAGGAACCATCCAGGCCACTCTGTAGTCTATCCAGGACTCAGACA 331		
QY	1770	CCGGATCCCCACCTGTCTCAGGAGCAGCCCTGACACTCTCTTTCAGACCTCTATTCCTTC 1829		
DB	320	CCGGATCCCCACCTGTCTCAGGAGCAGCCCTGACACTCTCTTTCAGACCTCTATTCCTTC 261		
QY	1830	CCAGAGATGTTGAGATGTTTCATCTCCAGCCCTGACCCCATGTCCTTGGACTCAGG 1889		
DB	260	CCAGAGATGTTGAGATGTTTCATCTCCAGCCCTGACCCCATGTCCTTGGACTCAGG 201		
QY	1890	GTCGTCTCCCGACATTTGGGTGACCGGTCTCTCTAGTTGAACCTGGGAACAATTC 1949		
DB	200	GTCGTCTCCCGACATTTGGGTGACCGGTCTCTCTAGTTGAACCTGGGAACAATTC 141		
QY	1950	CAAACTGTCTCAGGGGGGGTGTGGTCTCAATCTCTCTCCCTGGGCACTTTCATCTCAAGC 2009		

```
|||||
140 CAAAACGTCTCAGGGGGGGTTCGCTCAATCTCCTGGGCACTTTCATCTCAAGC 81
|||||
2010 TCAGGGCCCATCCTTCTCTGCGCTCTGACCAATTTAGTCCAGAAATAAATCTGAGA 2069
|||||
80 TCAGGGCCCATCCTTCTCTGCGCTCTGACCAATTTAGTCCAGAAATAAATCTGAGA 21
|||||
2070 AG 2071
||
20 AG 19

RESULT 2
US-09-996-243-308
: Sequence 308, Application US/09996243
: Patent No. 6478825
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi J.
: APPLICANT: Baker, Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Fong, Sherman
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2730P1C13
: CURRENT APPLICATION NUMBER: US/09/996,243
: CURRENT FILING DATE: 2001-11-14
: PRIOR APPLICATION NUMBER: 60/049787
: PRIOR FILING DATE: 1997-06-16
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/065186
: PRIOR FILING DATE: 1997-11-12
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066770
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/075945
: PRIOR FILING DATE: 1998-02-25
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/083322
: PRIOR FILING DATE: 1998-04-28
: PRIOR APPLICATION NUMBER: 60/084600
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/087106
: PRIOR FILING DATE: 1998-05-28
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087609
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087759
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087827
: PRIOR FILING DATE: 1998-06-19
: PRIOR APPLICATION NUMBER: 60/088021
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088025
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088026
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088028
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088029
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088030
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088033
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088326
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088167
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088202
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088212
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088217
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088655
: PRIOR FILING DATE: 1998-06-09
: PRIOR APPLICATION NUMBER: 60/088734
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088738
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088742
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088810
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088824
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088826
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088858
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/088861
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/088876
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/089105
: PRIOR FILING DATE: 1998-06-12
: PRIOR APPLICATION NUMBER: 60/089440
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089512
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089514
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089532
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089538
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089598
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089599
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089600
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089653
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/089907
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/089908
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/089947
: PRIOR FILING DATE: 1998-06-19
```

;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 23.3%; Score 482; DB 4; Length 1570;  
Best Local Similarity 100.0%; Pred. No. 8e-92;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGATTTGGGGGGCCTGTGGTCTGCAATGGCTCCCTCAGGGACTCGTGTCTCTGG 1649  
|||||  
DB 1077 CAGGGTGATTTGGGGGGCCTGTGGTCTGCAATGGCTCCCTCAGGGACTCGTGTCTCTGG 1136  
|||||  
QY 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTCTACACGAACTCTGCAAGTTC 1709  
|||||  
DB 1137 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTCTACACGAACTCTGCAAGTTC 1196  
|||||  
QY 1710 ACCAAGTGATCCAGGAAACCATCCAGGCAACTCTCTAGTATCCAGGACTCAGCACA 1769  
|||||  
DB 1197 ACCAAGTGATCCAGGAAACCATCCAGGCAACTCTCTAGTATCCAGGACTCAGCACA 1256  
|||||  
QY 1770 CCGGCATCCCACTGCTCAGGGACAGCCCTGACACTCTCTTCAGACCCCTCATTCCTTC 1829  
|||||  
DB 1257 CCGGCATCCCACTGCTCAGGGACAGCCCTGACACTCTCTTCAGACCCCTCATTCCTTC 1316  
|||||  
QY 1830 CCAGAGATTTGAGATGTTTCATCTCCAGCCCTGACCCCATGTCCTCTGGACTCAGG 1889  
|||||  
DB 1317 CCAGAGATTTGAGATGTTTCATCTCCAGCCCTGACCCCATGTCCTCTGGACTCAGG 1376  
|||||  
QY 1890 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGGGAACAATTC 1949  
|||||  
DB 1377 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGGGAACAATTC 1436  
|||||  
QY 1950 CAAAAGTGTCCAGGGCGGGGTTGGCTCTCAATCTCCCTGGGCACTTTTCATCCTCAAGC 2009  
|||||  
DB 1437 CAAAAGTGTCCAGGGCGGGGTTGGCTCTCAATCTCCCTGGGCACTTTTCATCCTCAAGC 1496  
|||||  
QY 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCCAAGAAATAAAGTGA 2069  
|||||  
DB 1497 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCCAAGAAATAAAGTGA 1556  
|||||  
QY 2070 AG 2071  
||  
DB 1557 AG 1558

RESULT 3  
US-08-824-874-2  
; Sequence 2, Application US/08824874  
; Patent No. 5962300  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,874  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:

## SEQUENCE CHARACTERISTICS:

LENGTH: 1476 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: KERANOT02  
CLONE: 820694  
US-08-824-874-2

Query Match 23.0%; Score 476; DB 2; Length 1476;

Best Local Similarity 100.0%; Pred. No. 1.4e-90;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGATTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGAGCTCGTGCTCTGG 1649  
DB 1001 CAGGGTGATTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGAGCTCGTGCTCTGG 1060  
QY 1650 GGAGATTACCTTGTGCGCGCGCCCAACAGACCGGGGTCTTACAGCAACCTCTTCAAGTTTC 1709  
DB 1061 GGAGATTACCTTGTGCGCGCGCCCAACAGACCGGGGTCTTACAGCAACCTCTTCAAGTTTC 1120  
QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCCTGAGTCAATCCAGGAGCTCAGCACA 1769  
DB 1121 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCCTGAGTCAATCCAGGAGCTCAGCACA 1180  
QY 1770 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1829  
DB 1181 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1240  
QY 1830 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCAATGCTCTCGGACTCAGG 1889  
DB 1241 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCAATGCTCTCGGACTCAGG 1300  
QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTGAACCTTGTGAGCAATTTTC 1949  
DB 1301 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTGAACCTTGTGAGCAATTTTC 1360  
QY 1950 CAAACTGTCCAGGGCGGGGTTGCGGTCTCAATCTCCCTGGGAGCTTTTCATCTCAAGC 2009  
DB 1361 CAAACTGTCCAGGGCGGGGTTGCGGTCTCAATCTCCCTGGGAGCTTTTCATCTCAAGC 1420  
QY 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAT 2065  
DB 1421 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAT 1476

## RESULT 4

US-09-210-084-2

Sequence 2, Application US/09210084

Patent No. 6197511

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Lal, Preeti

TITLE OF INVENTION: NOVEL KALLIKREIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/210,084

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/824,874

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0252 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1476 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: KERANOT02

CLONE: 820694

US-09-210-084-2

Query Match 23.0%; Score 476; DB 3; Length 1476;

Best Local Similarity 100.0%; Pred. No. 1.4e-90;

Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGATTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGAGCTCGTGCTCTGG 1649  
DB 1001 CAGGGTGATTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGAGCTCGTGCTCTGG 1060  
QY 1650 GGAGATTACCTTGTGCGCGCGCCCAACAGACCGGGGTCTTACAGCAACCTCTTCAAGTTTC 1709  
DB 1061 GGAGATTACCTTGTGCGCGCGCCCAACAGACCGGGGTCTTACAGCAACCTCTTCAAGTTTC 1120  
QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCCTGAGTCAATCCAGGAGCTCAGCACA 1769  
DB 1121 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCCTGAGTCAATCCAGGAGCTCAGCACA 1180  
QY 1770 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1829  
DB 1181 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1240  
QY 1830 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCAATGCTCTCGGACTCAGG 1889  
DB 1241 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCAATGCTCTCGGACTCAGG 1300  
QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTGAACCTTGTGAGCAATTTTC 1949  
DB 1301 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTGAACCTTGTGAGCAATTTTC 1360  
QY 1950 CAAACTGTCCAGGGCGGGGTTGCGGTCTCAATCTCCCTGGGAGCTTTTCATCTCAAGC 2009  
DB 1361 CAAACTGTCCAGGGCGGGGTTGCGGTCTCAATCTCCCTGGGAGCTTTTCATCTCAAGC 1420  
QY 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAT 2065  
DB 1421 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAT 1476

## RESULT 5

US-09-764-762-2

Sequence 2, Application US/09764762

Patent No. 6472195

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Lal, Preeti

TITLE OF INVENTION: NOVEL KALLIKREIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304



```
/
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA: US/09/764,762
/ APPLICATION NUMBER: US/09/764,762
/ FILING DATE: 16-Jan-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/210,084
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0252 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1476 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: KERANOT02
/ CLONE: 820694
/ SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-764-762-2

Query Match 23.0%; Score 476; DB 4; Length 1476;
Best Local Similarity 100.0%; Pred. No. 1.4e-90;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGATTCGGGGGGCTGTGCTGCTCAATGGCTCCCTCAGGGACTCGTGCTCTGG 1649
Db 1001 CAGGGTGATTCGGGGGGCTGTGCTGCTCAATGGCTCCCTCAGGGACTCGTGCTCTGG 1060
QY 1650 GGAGATTACCTTGTGCGGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1709
Db 1061 GGAGATTACCTTGTGCGGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1120
QY 1710 ACCAAGTGATCCAGGAACACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 1769
Db 1121 ACCAAGTGATCCAGGAACACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 1180
QY 1770 CCGGATCCCACTGCTGCAGGGACAGCCCTGACACTCTTTCAGACCCCTCATTCCTTC 1829
Db 1181 CCGGATCCCACTGCTGCAGGGACAGCCCTGACACTCTTTCAGACCCCTCATTCCTTC 1240
QY 1830 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1889
Db 1241 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1300
QY 1890 GTCTGCTCCCACTGAGGGTGCAGCGGTCTCTCTAGTTGAACCCCTGGGAACAATTC 1949
Db 1301 GTCTGCTCCCACTGAGGGTGCAGCGGTCTCTCTAGTTGAACCCCTGGGAACAATTC 1360
QY 1950 CAAAAGTTCAGGGGGGGGGTGGTCTCAATCTCCCTGGGGGCACTTTCATCCTCAAGC 2009
Db 1361 CAAAAGTTCAGGGGGGGGGTGGTCTCAATCTCCCTGGGGGCACTTTCATCCTCAAGC 1420
QY 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAT 2065
Db 1421 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAT 1476

RESULT 6
US-09-280-116-1
; Sequence 1, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: trypsin-like serine proteases
US-09-280-116-1

Query Match 22.2%; Score 460; DB 4; Length 1504;
Best Local Similarity 99.6%; Pred. No. 3.1e-87;
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1590 CAGGGTGATTCGGGGGGCTGTGCTGCTCAATGGCTCCCTCAGGGACTCGTGCTCTGG 1648
Db 1015 CAGGGTGATTCGGGGGGCTGTGCTGCTCAATGGCTCCCTCAGGGACTCGTGCTCTGG 1074
QY 1649 GGGAGATTACCTTGTGCGGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTT 1708
Db 1075 GGGAGATTACCTTGTGCGGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTT 1134
QY 1709 CACCAAGTGATCCAGGAACACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGC 1768
Db 1135 CACCAAGTGATCCAGGAACACCATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGC 1194
QY 1769 ACCGGCATCCCACTGCTGCAGGGACAGCCCTGACACTCTTTCAGACCCCTCATTCCTT 1828
Db 1195 ACCGGCATCCCACTGCTGCAGGGACAGCCCTGACACTCTTTCAGACCCCTCATTCCTT 1254
QY 1829 CCCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAG 1888
Db 1255 CCCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAG 1314
QY 1889 GGTCTGCTTCCCACTGCTGCAGGGTGGCTGCTCTCTAGTTGAACCCCTGGGAACAATTT 1948
Db 1315 GGTCTGCTTCCCACTGCTGCAGGGTGGCTGCTCTCTAGTTGAACCCCTGGGAACAATTT 1374
QY 1949 CCAAAAGTCTCCAGGGCGGGGTGGTGTCTCAATCTCCCTGGGGCACTTTTCATCCTCAAG 2008
Db 1375 CCAAAAGTCTCCAGGGCGGGGTGGTGTCTCAATCTCCCTGGGGCACTTTTCATCCTCAAG 1434
QY 2009 CTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGT-CCCAGAAATAAATGA 2067
Db 1435 CTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGT-CCCAGAAATAAATGA 1494
QY 2068 GAAG 2071
Db 1495 GAAG 1498

RESULT 7
US-09-128-155-17
; Sequence 17, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match      10.7%; Score 222.6; DB 3; Length 176373;
Best Local Similarity 84.0%; Pred. No. 2.9e-37;
Matches 263; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 1058 AAAGGTGAACCTTGGAGGCGCAGCATGGTGGCTCAGCCCTGTATCCCAACACTTT-GGA 1116
Db 127012 AAGAAAAAATTTATTTGGCCAGCAGCGCGCTCATGCCCTATATCCAGCACTTTGGGA 127071

QY 1117 GGCTGAGGTGGGCAATCACTTGAGGCCAGAGTTCGAGACCAGCCTGGCCAAACATGGTG 1176
Db 127072 GGCTGAGGCAAGTGGATCACTGAGGTTCAGAGACCAAGCTGGCCAAACATGGTG 127131

QY 1177 AAACCCCGTCTCTACAAAAAATACAAAAAATAGCCGGGTGTGTGATGGACACTGT 1236
Db 127132 AAACCCCATCTCTACTAAAAAATACAAAAAATAGCCAGGTGTGTGATGGACACTGT 127191

QY 1237 AGTCACAGCTACTTGGAGGCTGAGGCGAGAGATTCCTTGAACCCGGGAGATGGAGGCT 1296
Db 127192 ANTCCAGCTACACGGGAGGCTGAGGCGAGAGATCGCTTGAACCCGGGAGGAGGAT 127251

QY 1297 GCAGTGAGCTGAGTCAAGGCACTGCGCTCCACCTGGGCAACAGAGTAAGACTCCACT 1356
Db 127252 GCAGTGAGCAAGTTCAAGCACTGACCCAGGCTGAGGCGAGAGATCGCTTGAACCCGGGAGGAGGAT 127311

QY 1357 CAAAAAARAAA 1369
Db 127312 CAAAAAARAAA 127324
```

```
RESULT 8
US-09-851-896-3/C
; Sequence 3, Application US/09851896
; Patent No. 6410325
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT)
; FILE REFERENCE: RTS-0220
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 70000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-851-896-3

Query Match      10.6%; Score 219.4; DB 4; Length 70000;
Best Local Similarity 85.9%; Pred. No. 1.1e-36;
Matches 255; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 1074 GGCAGGCAATGGTGGCTCAGCCCTGTATCCCAACACTTT-GGAGGCTGAGGTGGGCGAA 1132
Db 36372 GGCAGGCAATGGTGGCTCAGCCCTGTATCCCAACACTTTGGGAGGCGCAAGGAGGCGAGA 36313

QY 1133 TCACCTGAGGCGCAGGAGTTCGAGACCAAGCTGGCCAAACATGGTGAACCCCGCTCTTACA 1192
Db 36312 TCACCTGAGGTAGGAGTTCGAGACCAAGCTGGCCAAACATGGTGAACCCCGCTCTTACT 36253

QY 1193 AAAAAAATACAAAAAATAGCCGGGTGTGTGATGGACACTGTAGTCACAGCTACTGTG 1252
```

```
Db 36252 AAAAAACACAAAAAATAGCCGGGTGTGTGCGGCGCTAGTCCAGCTACGCGG 36193
QY 1253 GAGGCTGAGGCGAGGAGAAATTTGAACCCGGGAGATGGAGGCTGCAGTGAGCTGAGGTC 1312
Db 36192 GAGGCTGAGGCGAGGAGAAATTTGAACCCGGGAGGAGCTTGCAGTGAGGCGGAGATC 36133
QY 1313 AGGCCACTGCGTCCCAACCTGGCCAAACAGAGTAAGACTTCCATCTCAAAAAA 1369
Db 36132 GCGCCACTGCATCCAGCCTGGCGCAGACAGAGCGAGACTCGCTCTCAAAAAA 36076

RESULT 9
US-09-128-155-16
; Sequence 16, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(152331)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

Query Match      10.6%; Score 219.2; DB 3; Length 152331;
Best Local Similarity 81.4%; Pred. No. 1.4e-36;
Matches 254; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1058 AAAGGTGAACCTTGGAGGCGCAGCATGGTGGCTCAGCCCTGTATCCCAACACTTTGGAG 1117
Db 58133 AAGAAAAAATTTATTTGGCCAGGCGAGGCTCATGCCCTATAATCCAGCACTTTGGGA 58192

QY 1118 GCTGAGGTGGCGAATCACTTTGAGGCGCAGAGTTCGAGACCGCCTGGCCAAACATGGTGA 1177
Db 58193 GGCTGAGGCGAGGTGATCACTGAGGTTCGAGAGTTCGAGACCGCCTGGCCAAACATGGTGA 58252

QY 1178 AACCCGCTCTCTACAAAAAATACAAAAAATAGCCGGGTGTGTGATGGACACCTGTGA 1237
Db 58253 AACCCCATCTCTACTAAAAAATACAAAAAATAGCCAGGTGTGTGTTGGGCGACCTGTGA 58312

QY 1238 GTCACAGCTACTTTGGGAGGCTGAGGCGAGAGAAATTCCTTGAACCCGGGAGATGGAGCTG 1297
Db 58313 ATCCAGCTACACGGGAGGCTGAGGCGAGGAGAAATTCGTTCAACCCGGGAGGAGAAATG 58372

QY 1298 CAGTGAGCTGAGTTCAGGCGCACTGCGCTCCAACTGGGCAACAGAGTAAGACTCCATCTC 1357
Db 58373 CAGTGAGCGCAAGTTCAGGCGCACTGCACCCCGCCTGGGCAACAGAGCAAGACTTTGTC 58432

QY 1358 AAAAAAARAAA 1369
Db 58433 CAAAAAARAAA 58444
```

```
RESULT 10
US-09-326-480A-4
; Sequence 4, Application US/09326480A
; Patent No. 6551792
; GENERAL INFORMATION:
```

APPLICANT: Marta Blumenfeld  
APPLICANT: Ilija Tchoumakov  
APPLICANT: Henri-Jean Garchon  
APPLICANT: Jean-Francois Bach  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE WHICH ENCODES A  
TITLE OF INVENTION: FLAVIN MONOOXYGENASE, THE CORRESPONDING PROTEIN AND THEIR  
FILE REFERENCE: GENSSET.064C1  
CURRENT FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: PCT/FR97/02226  
PRIOR FILING DATE: 1997-12-05  
PRIOR APPLICATION NUMBER: FR 96/15032  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 25464  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE: UNSURE  
OTHER INFORMATION: genomic  
US-09-326-480A-4

Query Match 10.6%; Score 218.6; DB 4; Length 25464;  
Best Local Similarity 85.0%; Pred. No. 1.3e-36;  
Matches 256; Conservative 0; Mismatches 44; Indels 1; Gaps 1;  
QY 1070 TGGAGCCAGGATGTGTGCTCAGCGCTGTAATCCCAACACTTTG-GAGGCTGAGGTGG 1128  
DB 18484 TGCAGATTGGGATGTGTGCTCAGCGCTGTAATCCCAACACTTTGTTGAGGCTGAGGCGG 18543  
QY 1129 CGAATCACTTGGAGCCAGGATTCGAGACGAGCTGCGCAACATGTTGAACCCCGCTCTC 1188  
DB 18544 CGGATTACCTGAGTTCGAGAGTTCGAGACGAGTCTGCGCAACATGTTGAACCCCGCTCTC 18603  
QY 1189 TACAAAAAATACAAAAAATAGCCGGTGTGTGATGGACACCTGTAGTCACAGCTAC 1248  
DB 18604 TACTAAAAATACAAAAAATAGTCAGTGTGTGCTGACCTGTATCCAGCTAC 18663  
QY 1249 TTGGGAGGCTGAGGAGGAGATTCCTTGAACCCGGGAGATGGAGCTGAGTGTGAGCTGA 1308  
DB 18664 TCGGGAGGCTGAGGAGGAGATTCCTTGAATCAGGAGGAGTGTGAGGCTGAGGCTGAGCCAA 18723  
QY 1309 GGTACGCGCACTGCGCTCCAACTGGGCAACAGAGTAAGACTCCATCTCAAAAAA 1368  
DB 18724 GATCGCGCACTGCACTAGCTAGCTGACAGAGTGAAGTCCATCTCAAAAAA 18783  
QY 1369 A 1369  
DB 18784 A 18784

RESULT 11  
US-09-404-879A-1/c  
Sequence 1, Application US/09404879A  
Patent No. 6468546  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Algate, Paul A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462C2  
CURRENT APPLICATION NUMBER: US/09/404,879A  
CURRENT FILING DATE: 1999-09-24  
NUMBER OF SEQ ID NOS: 393  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 461  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-404-879A-1

Query Match 10.4%; Score 216.4; DB 4; Length 461;  
Best Local Similarity 84.4%; Pred. No. 1.5e-36;  
Matches 255; Conservative 0; Mismatches 46; Indels 1; Gaps 1;  
QY 1069 TTGGAGCCAGGATGTGTGCTCAGCGCTGTAATCCCAACACTTTT-GGAGGCTGAGGTGG 1127  
DB 368 TTGGGCGCGGCGTGTGTGCTCAGCGCTGTAATCCCAACACTTTTGGGAGGCGGAGCGG 309  
QY 1128 GCGAATCACTTGGAGCCAGGAGTTCGAGACGAGCTTGGCCCAACATGGTGAACCCCGTCT 1187  
DB 308 GTGGATCACTTGGAGTCAAGAGAGCTTGAAGAGAGCTTGGCCCAACCTGGTGAACCCCGTCT 249  
QY 1188 CTACAAAAAATACAAAAAATAGCCGGTGTGTGATGGACACCTGTAGTCACAGCTA 1247  
DB 248 CTACTAAAAATACAAAAAATAGCTGAGCGTGTGTGCGGGCGCTGTAATCCAGCTA 189  
QY 1248 CTTGGAGGCTGAGGAGGAGATTCCTTGAACCCGGGAGATGGAGGCTGAGCTGAGCTG 1307  
DB 188 CTTGGAGGCTGAGGAGGAGATTCCTTGAACCTGGAGGCGGAGGTTGCAGCGAGCTG 129  
QY 1308 AGTCAAGGCTGAGGCTGAGGAGGAGTTCGAGACGAGTGAAGACTCCATCTCAAAAAA 1367  
DB 128 AGATCATGCGTGTGTGCTCAGCTTGGGCAACAGAGTGAAGTCCATCTCAAAAAA 69  
QY 1368 AA 1369  
DB 68 AA 67

RESULT 12  
US-09-404-879A-3/c  
Sequence 3, Application US/09404879A  
Patent No. 6468546  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Algate, Paul A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462C2  
CURRENT APPLICATION NUMBER: US/09/404,879A  
CURRENT FILING DATE: 1999-09-24  
NUMBER OF SEQ ID NOS: 393  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 461  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-404-879A-3

Query Match 10.4%; Score 216.4; DB 4; Length 461;  
Best Local Similarity 84.4%; Pred. No. 1.5e-36;  
Matches 255; Conservative 0; Mismatches 46; Indels 1; Gaps 1;  
QY 1069 TTGGAGCCAGGATGTGTGCTCAGCGCTGTAATCCCAACACTTTT-GGAGGCTGAGGTGG 1127  
DB 368 TTGGGCGCGGCGTGTGTGCTCAGCGCTGTAATCCCAACACTTTTGGGAGGCGGAGCGG 309  
QY 1128 GCGAATCACTTGGAGCCAGGAGTTCGAGACGAGCTTGGCCCAACATGGTGAACCCCGTCT 1187  
DB 308 GTGGATCACTTGGAGTCAAGAGAGCTTGAAGAGAGCTTGGCCCAACCTGGTGAACCCCGTCT 249  
QY 1188 CTACAAAAAATACAAAAAATAGCCGGTGTGTGATGGACACCTGTAGTCACAGCTA 1247  
DB 248 CTACTAAAAATACAAAAAATAGCTGAGCGTGTGTGCGGGCGCTGTAATCCAGCTA 189  
QY 1248 CTTGGAGGCTGAGGAGGAGATTCCTTGAACCCGGGAGATGGAGGCTGAGCTGAGCTG 1307  
DB 188 CTTGGAGGCTGAGGAGGAGATTCCTTGAACCTGGAGGCGGAGGTTGCAGCGAGCTG 129  
QY 1308 AGTCAAGGCTGAGGCTGAGGAGGAGTTCGAGACGAGTGAAGACTCCATCTCAAAAAA 1367  
DB 128 AGATCATGCGTGTGTGCTCAGCTTGGGCAACAGAGTGAAGTCCATCTCAAAAAA 69

```
Db 128 AGATCATGCCGTTGTACTCCAGCTTGGGCAACAGAGTGAGACTCCATCTCAAAACAACAAAC 69
QY 1368 AA 1369
Db 68 AA 67

RESULT 13
US-09-338-933-1/c
; Sequence 1, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-1

Query Match 10.4%; Score 216.4; DB 4; Length 461;
Best Local Similarity 84.4%; Pred. No. 1.5e-36;
Matches 255; Conservative 0; Mismatches 46; Indels 1; Gaps 1;

QY 1069 TTGGAGGCCAGGATGGTGTGCTCACGCCCTGTATATCCCAACACTTTT-GGAGGCTGAGGTGG 1127
Db 368 TTGGGGCGGGCGTGTGTGCTCACGCCCTGTATATCCCAACACTTTTGGGAGCGCGAGCGG 309
QY 1128 GCGAATCACTTGAGGCCAGGAGTTCCAGACACCGCTGGCCAAACATGTTGAAACCCCGTCT 1187
Db 308 GTGGATCACTTGAGGCTGAGGAGTTCAAGAGACGCCCTGGCCAAACCTGTTGAAACCCCGTCT 249
QY 1188 CTACAAAAAATACAAAAAATTAGCCGGGTGTGTTGATGGACACCTGTAGTCACAGCTA 1247
Db 248 CTACTAAAAATACAAAAAATTAGCTGAGCGTGTGGCGGCCCTGTATATCCCACTA 189
QY 1248 CTTGGGAGGCTGAGGAGGAGAAATGCTTGAACCCCGGAGATGGAGCTGCAGTGAGCTG 1307
Db 188 CTTGGGAGGCTGAGGAGGAGAAATCACTTGAACGTGGGAGCGGAGGTTGCAGCGAGCTG 129
QY 1308 AGGTCAGGCCACTCGCGCTCCAACTGGGCAACAGAGTAAGACTCCATCTCAAAAAA 1367
Db 128 AGATCATGCCGTTGTACTCCAGCTTGGGCAACAGAGTGAGACTCCATCTCAAAACAAC 69
QY 1368 AA 1369
Db 68 AA 67

RESULT 14
US-09-338-933-3/c
; Sequence 3, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-3

Query Match 10.4%; Score 216.4; DB 4; Length 461;
Best Local Similarity 84.4%; Pred. No. 1.5e-36;
Matches 255; Conservative 0; Mismatches 46; Indels 1; Gaps 1;

QY 1069 TTGGAGGCCAGGATGGTGTGCTCACGCCCTGTATATCCCAACACTTTT-GGAGGCTGAGGTGG 1127
Db 368 TTGGGGCGGGCGTGTGTGCTCACGCCCTGTATATCCCAACACTTTTGGGAGCGCGAGCGG 309
QY 1128 GCGAATCACTTGAGGCCAGGAGTTCCAGACACCGCTGGCCAAACATGTTGAAACCCCGTCT 1187
Db 308 GTGGATCACTTGAGGCTGAGGAGTTCAAGAGACGCCCTGGCCAAACCTGTTGAAACCCCGTCT 249
QY 1188 CTACAAAAAATACAAAAAATTAGCCGGGTGTGTTGATGGACACCTGTAGTCACAGCTA 1247
Db 248 CTACTAAAAATACAAAAAATTAGCTGAGCGTGTGGCGGCCCTGTATATCCCACTA 189
QY 1248 CTTGGGAGGCTGAGGAGGAGAAATGCTTGAACCCCGGAGATGGAGCTGCAGTGAGCTG 1307
Db 188 CTTGGGAGGCTGAGGAGGAGAAATCACTTGAACGTGGGAGCGGAGGTTGCAGCGAGCTG 129
QY 1308 AGGTCAGGCCACTCGCGCTCCAACTGGGCAACAGAGTAAGACTCCATCTCAAAAAA 1367
Db 128 AGATCATGCCGTTGTACTCCAGCTTGGGCAACAGAGTGAGACTCCATCTCAAAACAAC 69
QY 1368 AA 1369
Db 68 AA 67

RESULT 15
US-09-215-681-1/c
; Sequence 1, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF OVARIAN CANCER
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-1

Query Match 10.4%; Score 216.4; DB 4; Length 461;
Best Local Similarity 84.4%; Pred. No. 1.5e-36;
Matches 255; Conservative 0; Mismatches 46; Indels 1; Gaps 1;

QY 1069 TTGGAGGCCAGGATGGTGTGCTCACGCCCTGTATATCCCAACACTTTT-GGAGGCTGAGGTGG 1127
Db 368 TTGGGGCGGGCGTGTGTGCTCACGCCCTGTATATCCCAACACTTTTGGGAGCGCGAGCGG 309
QY 1128 GCGAATCACTTGAGGCCAGGAGTTCCAGACACCGCTGGCCAAACATGTTGAAACCCCGTCT 1187
Db 308 GTGGATCACTTGAGGCTGAGGAGTTCAAGAGACGCCCTGGCCAAACCTGTTGAAACCCCGTCT 249
QY 1188 CTACAAAAAATACAAAAAATTAGCCGGGTGTGTTGATGGACACCTGTAGTCACAGCTA 1247
Db 248 CTACTAAAAATACAAAAAATTAGCTGAGCGTGTGGCGGCCCTGTATATCCCACTA 189
QY 1248 CTTGGGAGGCTGAGGAGGAGAAATGCTTGAACCCCGGAGATGGAGCTGCAGTGAGCTG 1307
Db 188 CTTGGGAGGCTGAGGAGGAGAAATCACTTGAACGTGGGAGCGGAGGTTGCAGCGAGCTG 129
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 5, 2003, 09:57:15 ; Search time 3128 Seconds  
(without alignments)  
3832.008 Million cell updates/sec

Title: US-09-936-271B-14

Perfect score: 1608

Sequence: 1 MATARPPWVLCALITALL.....VYTNLCKFTKWIQTQANS 293

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n model -DEV=xlp  
-Q/Cgn2\_1/USPTO.spool/US09936271/runat\_04082003\_104245\_25702/app\_query.fasta\_1.455  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09936271 @CGN\_1\_1.2326 @runat\_04082003\_104245\_25702 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pln.\*  
35: em.htg\_rod.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em.sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1608	100.0	882	9	BT006867
2	1608	100.0	882	12	BT007831
3	1608	100.0	1370	9	AY279380 Homo sapi
4	1608	100.0	1381	6	BD107879 36 human
5	1608	100.0	1387	9	AF168768 Homo sapi
6	1608	100.0	1438	9	AY279381 Homo sapi
7	1608	100.0	1499	6	BD005362 Protease.
8	1608	100.0	1516	6	BD107865 36 human
9	1608	100.0	1536	9	BC008036 Homo sapi
10	1608	100.0	1570	6	AR252580 Sequence
11	1608	100.0	1570	6	AX080829 Sequence
12	1608	100.0	1570	6	AX03421 Sequence
13	1608	100.0	1570	6	AX464322 Sequence
14	1595	99.2	1476	6	AR078184 Sequence
15	1595	99.2	1476	6	AR137506 Sequence
16	1595	99.2	1476	6	AR242358 Sequence
17	1595	99.2	1476	6	BD082659 Keratinoc
18	1533	94.1	1439	6	BD107880 36 human
19	1475	91.7	929	9	AF435981 Homo sapi
20	1475	91.7	1132	9	AF435980 Homo sapi
21	1403	87.3	1504	6	AR263823 Sequence
22	1002	62.3	11570	9	AF135028 Homo sapi
23	1002	62.3	217346	2	AC027602 Homo sapi
24	1002	62.3	230000	9	AF243527 Homo sapi
25	997	62.0	107487	9	AC011483 Homo sapi
26	985.5	61.3	200792	2	AC130782 Pan trogl
27	725	45.1	735	6	AX429955 Sequence
28	725	45.1	738	6	BD139877 Compounds
29	688	42.8	673	6	AX067353 Sequence
30	676	42.0	1140	4	U76256 Sus scrofa
31	671	41.7	765	6	AR261054 Sequence
32	671	41.7	765	6	AR278585 Sequence
33	671	41.7	765	6	AX141034 Sequence
34	671	41.7	765	6	AX200894 Sequence
35	671	41.7	765	6	AX267550 Sequence
36	671	41.7	765	9	AF259969 Homo sapi
37	670	41.7	1347	9	AF113140 Homo sapi
38	654.5	40.7	1052	6	AR219287 Sequence
39	652.5	40.6	1166	6	AR152173 Sequence
40	650.5	40.5	833	6	AR060847 Sequence
41	650.5	40.5	833	6	BD082136 Novel pro
42	650.5	40.5	1106	6	AX016287 Sequence
43	650.5	40.5	1106	6	BD137019 Human CAS
44	650.5	40.5	1158	6	AX016289 Sequence
45	650.5	40.5	1158	6	BD137020 Human CAS

#### ALIGNMENTS

RESULT 1

BT006867  
LOCUS 882 bp mRNA linear PRI 13-MAY-2003  
DEFINITION Homo sapiens kallikrein 5 mRNA, complete cds.  
ACCESSION BT006867  
VERSION BT006867.1 GI:30582572  
KEYWORDS FLI\_CDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 882)  
AUTHORS Kalline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,  
Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y.,  
Phelan, M., and Farmer, A.  
TITLE Cloning of human full-length CDSs in BD Creator(TM) System Donor  
vector  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 882)  
AUTHORS Kalline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,  
Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y.,  
Phelan, M., and Farmer, A.  
TITLE Direct Submission  
JOURNAL Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow  
Circle, Palo Alto, CA 94303, USA  
COMMENT This CDS clone is a part of a collection of human full length  
expression clones generated by BD Biosciences Clontech and the  
Harvard Institute of Proteomics. Each CDS has been cloned in two  
forms: with and without stop-codon (to allow fusion with C-terminal  
tag). The CDS has been directionally cloned using BD In-Fusion(TM)  
cloning system between the Sali and HindIII sites of the pDNR-DUAL  
vector. Additional sequences in the clone: 'ACC' after Sali site  
and before 'ATG' to provide Kozak consensus sequence; 'GG' after  
last codon and before HindIII site to maintain reading frame.  
Clone distribution: <http://bioinfo.clontech.com/orfclones>.  
FEATURES  
source  
1..882  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GH00284X1.0"  
/clone\_lib="BD Creator(TM) CDS Library derived from MGC  
collection"  
/lab\_host="DH5alpha T1 resistant"  
/note="vector: pDNR-Dual"  
1..882  
/codon\_start=1  
/product="kallikrein 5"  
/protein\_id="AAP35513.1"  
/db\_xref="GI:30582573"  
/translation="MATARPMMVLCALITALLGVTEHVHVLANNVSDHPNTPVS  
GSNODLGAGEDARSDDSSRIINGSDCMHTQWPQAAALLRPNOLYCGAVLHPQM  
LTAACHKRVFRLGHYSLSPIYESGQMFQGVKSIPIHPGYSHPGHNDLMLIKLN  
RRIRPTKDVPRINVSHPKAGTKCLVSGWGTKSPQVHPFKVLQCLNISVLSQKRC  
DAYPROIDDPMFCAGDKAGRDSQGGSGPVCVNGSLQGLVSNQDYPFCARPNRPVYT  
NLCKFTWKIQTQANS"  
BASE COUNT 193 a 272 c 237 g 180 t

## Alignment Scores:

Pred. No.: 8,11e-121 Length: 882  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-936-271b-14 (1-293) x BT006867 (1-882)

Qy 1 MetalThrAlaArgProToTrpMetTrpValLeuCysAlaLeuLeuThrAlaLeuLeu 20  
|||||  
Db 1 ATGGCTACAGCAAGACCCCTGGTGTGGTGTCTGTCTGTATCATCAGACCCCTGCTT 60  
|||||  
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40

Db 61 CTGGGGGTACAGAGATGTTCTCGCAACAATGATTTTCTGTGTACCCCTCTAAC 120  
|||||  
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60  
|||||  
Db 121 ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCCGGGAGACGCCGCTG 180  
|||||  
Qy 61 AspAspSerSerArgIlelleAsnGlySerAspCysAspMetHisThrGlnProTrp 80  
|||||  
Db 181 GATGACAGCAGCAGCGCATCATATGATCGACTCGCATATGCACACCCAGCCGTGG 240  
|||||  
Qy 81 GlnAlaLeuLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
|||||  
Db 241 CAGCGCCGCTGTGTAAAGCCCAACACGCTTACTCGGGGGGGTGTGGTGCATCCA 300  
|||||  
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgGlySerValPheArgValArgLeuGlyHis 120  
|||||  
Db 301 CAGTGGCTGTACAGCGCCGCCACTCGAGCAAGAAAGTTTTCAGAGTCCGCTCTGGCCAC 360  
|||||  
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
|||||  
Db 361 TACTCCCTGTCAACAGTTTATGAATCTGGGCAGCAGATGTTCAGGGGGTCAATCATC 420  
|||||  
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160  
|||||  
Db 421 CCCACCCCTGGTACTCCACCCCTGGCCACTCTAACGACCTCATGCTCATCAAACTGAAC 480  
|||||  
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
|||||  
Db 481 AGAAGATTCTGCCCTCCACTAAAGATGTTCAGACCCATCAACGCTCTCTCTATTGTCCTCT 540  
|||||  
Qy 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
|||||  
Db 541 GCTGGGACAAAGTGTGTGTCTGTGGTGGGGGACCAAGAGCCGCCCAAGTGCACATTC 600  
|||||  
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
|||||  
Db 601 CCTAAGGTCTCTCCAGTCTTGAATATCAGCGTGTCTAAGTCAGAAAGTGGCAGATGCT 660  
|||||  
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
|||||  
Db 661 TACCGAGACAGATAGATGACACCATGTCTCGCCCGGTGCACAAAGCAGGTAGAGACTCC 720  
|||||  
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
|||||  
Db 721 TGGCAGGGTGATTTCTGGGGGGCTGTGTGTCTGCAATGGCTCCCTGCAGGGACTCTGTGTC 780  
|||||  
Qy 261 TrpGlyAspTyrProCysAlaAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
|||||  
Db 781 TGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAG 840  
|||||  
Qy 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
|||||  
Db 841 TTCACCAAGTGTGATCCAGGAAACCATCCAGGCCCAACTCC 879  
|||||

## RESULT 2

BT007831  
LOCUS 882 bp mRNA linear SYN 13-MAY-2003  
DEFINITION Synthetic construct Homo sapiens kallikrein 5 mRNA, partial cds.  
ACCESSION BT007831  
VERSION BT007831.1 GI:30584500  
KEYWORDS FLI\_CDNA.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 882)  
AUTHORS Kalline, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,  
Koundinya, M., Raphael, J., Moreira, D., Kelley, T., Labaer, J., Lin, Y.,  
Phelan, M., and Farmer, A.  
TITLE Cloning of human full-length CDSs in BD Creator(TM) System Donor  
vector  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 882)



**AUTHORS** Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S., Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y., Phelan,M. and Farmer,A.

**TITLE** Direct Submission

**JOURNAL** Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow Circle, Palo Alto, CA 94303, USA

**COMMENT** This CDS clone is a part of a collection of human full length expression clones generated by BD Biosciences Clontech and the Harvard Institute of Proteomics. Each CDS has been cloned in two forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the SalI and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after SalI site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://bioinfo.clontech.com/orfclones>.

# FEATURES

Location/Qualifiers  
1..882  
/organism="synthetic construct"  
/mol\_type="mRNA"  
/db\_xref="taxon:32630"  
/clone="GH00284L1.0"  
/clone\_lib="BD Creator(TM) CDS Library derived from MGC collection"  
/lab\_host="DH5alpha T1 resistant"  
/note="Vector: pDNR-Dual"  
1..->882  
/note="Mutations: 881:Stop->Leu"  
/codon\_start=1  
/transl\_table=1  
/product="Homo sapiens kallikrein 5"  
/protein\_id="AAP36503.1"  
/db\_xref="GI:30584501"  
/translation="NATARPMMWVLCALITALLGVTEHVLANNVSDHPSTNTPVS  
GSDNQLGAGEDARDSDSSRIINGSCDHTQWQAALLRLPNQLYCGAVLVHPQW  
LITAAHCKRVFRVRLGHYSUSPVYSGQMFQGVKSIPHPGYSHPGHNSDMLIKLN  
RRIRPKDRPVINYSHGCSAGTKCLVSGWGTTPKSPQVHPFKVLQCLNIVLSQRKE  
DAYPRQIDTFMFCAGDKAGKSCGDSGGPVVYVNGSLQGLVSWGDPYPCARPNRPVYT  
NLCKFTKWIQETIQANSL"

## CDS

BASE COUNT 192 a 272 c 237 g 181 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8,11e-121 Length: 882  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0  
US-09-936-271B-14 (1-293) x BT007831 (1-882)

QY 1 MetaAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
Db 1 ATGGGTACAGCAAGACCCCTGGATGTGGGTCTGTCTGTCTATCACACCTTGCTT 60  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
Db 61 CTGGGGGTACAGACATGTTCTCCCAACAATGATGTTCTGTGACCCCTCTATAC 120  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyCysAlaArgSer 60  
Db 121 ACCGTGCTCTCGGAGCAACACGAGCTGGGAGCTGGGGCGGGGAAGACCCCGGTG 180  
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80  
Db 181 GATGACAGCAGCGCATCATCAATGATCCGATCGGATGATGATGATGATGATGATG 240  
QY 81 GlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100  
Db 241 CAGGGCGGCTGTTCTTAAAGCCCCAACACGCTCTACTGCGGGCGGTGTGTGTGATCA 300  
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120

Db 301 CAGTGGCTGCTCAGCGCGCCCTCCACATGTCAGAGAAAGTTTTCAGAGTCTGCGCCAC 360  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheClnGlyValLysSerIle 140  
Db 361 TACTCCCTGTCCACAGTTTATGAATGCGGACAGATGTCAGGGGTCAATCATC 420  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
Db 421 CCCCACCTGGCTACTCCCACTGCGCACTCTAACGACCTCATGCTCATCAACTGAAC 480  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
Db 481 AGAAGAATTCGTCCTCCACTAAAGATGTCAGACCATCAAGCTCTCTCTCATGTCCTCT 540  
QY 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200  
Db 541 GCTGGACAAAGTGTGTGTGTGGTGGGGGACCAACCAAGAGCCCCCAAGTGCACTTC 600  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
Db 601 CCTAAGGTCTCCAGTGTCTGAATATCAGCGTCTAAGTCAGAAAAGTCCGAGGATGCT 660  
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
Db 661 TACCCGAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAAGCAGTAGAGACTCC 720  
QY 241 CysGlnGlyAspSerGlyClyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
Db 721 TGCCAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGACTCGTGTC 780  
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
Db 781 TGGGGAGATTACCTGTGTCGGCGGCCAACACACGCGGTGTCTACACGAACTCTGCAAG 840  
QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
Db 841 TTCACCAAGTGTGATCCAGGAAACCATCCAGGCAACTCC 879  
RESULT 3  
LOCUS AY279380 1370 bp mRNA linear PRI 26-MAY-2003  
DEFINITION Homo sapiens kallikrein 5 splice variant 1 (KLK5) mRNA, complete  
cvs; alternatively spliced.  
ACCESSION AY279380  
VERSION AY279380.1 GI:31075480  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1370)  
AUTHORS Kurlender,L., Yousef,G.M., White,N.M.A., Robb,J.-D., Borgono,C.A. and Diamandis,E.P.  
TITLE Identification of splice variants for the human kallikrein gene 5 (KLK5)  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 1370)  
Kurlender,L., Yousef,G.M., White,N.M.A., Robb,J.-D., Borgono,C.A. and Diamandis,E.P.  
TITLE Direct Submission  
JOURNAL Submitted (19-APR-2003) Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada  
FEATURES  
source Location/Qualifiers  
1..1370  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.4"  
1..1370  
/gene="KLK5"

```

5'UTR      1. .127
/gene="KLK5"
CDS        128. .1009
/gene="KLK5"
/notes="serine protease; hk5; alternatively spliced;
similar to the product of GenBank Accession Number
AF135028"
/codon_start=1
/product="kallikrein 5 splice variant 1"
/protein_id="AAP42275.1"
/db_xref="GI:31075481"
/translation="MATARPWWVLCALITALLLGVTEVLANNVSDHPNTVPS
GNSQDLGAGAGEDARDSDSSRLINGSDDMTQPWQALLLPNQLYCGAVLVHPQW
LTPAARCKKRVFRVLRGHYSLSPIYESGQMFQVKSIPHPGYSHPGHNDMLIKLN
RRIRPTKVRPIINVSHCPISAGTKCLVSGWGTTPQVHFPPKVLQCLNTSVLSQKRC
DAYPROIDDTMFCAGDKAGRDSQGDSPVVCNGSLQGLVSGWDYPCARPNRPGVYT
NLCKFTKWIQETIQANS"
1010. .1370
/gene="KLK5"
polyA_signal 1319. .1324
/gene="KLK5"
BASE COUNT 323 a 424 c 352 g 271 t
ORIGIN

Alignment Scores:
Pred. No.: 1.3e-120 Length: 1370
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-936-271B-14 (1-293) x AV279380 (1-1370)

Qy 1 MetAlaThrAlaArgProTriPmetTriPValLeuCysAlaLeuIleThrAlaLeu 20
Db 128 ATGGCTACACAGACCCCTGGATGGTGGTCTCTGCTCTGATCAGCGCTTGCTT 187

Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 188 CTGGGGGTACAGAGCATGTTCTCGCCCAACAATGATGTTTCTGTGACCACTCTAAC 247

Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 248 ACCGTGCCCCCTGGGAGCAACAGACCTGGGAGCTGGGCGGGGGAAGACGCCGCTCG 307

Qy 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTriP 80
Db 308 GATGACAGCAGACCGCATCATCATGATCGACTGCGATATGCACACCCAGCGCTGG 367

Qy 81 GlnAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 368 CAGGCCGCGTGTGCTGAAGGCCCAACAGCTCTACTCGGGCGGGTGTGGTGCATCCA 427

Qy 101 GlnTriPLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
Db 428 CAGTGGCTGTACGGCGGCCACTCGAGGAAGAAGTTTTCAGAGTCCGCTCGGCCAC 487

Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 488 TACTCCCTGTACACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGGTCAATCATC 547

Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
Db 548 CCCACCCCTGGCTACCTCCACCTGGCCACTCTAACGACCTCATGCTCATCAACTGAAC 607

Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 608 AGAAGATTCTGCTCCCACTAAAGATGTGAGACCCATCAACGCTCTCTCTATTGTCCTCT 667

Qy 181 AlaGlyThrLysCysLeuValSerGlyTriPLeuThrLysSerProGlnValHisPhe 200
Db 668 GCTGGGCAAGTGTGGTGTCTGGCTGGGGGACCAACCAAGAGCCCCCAAGTGCACCTTC 727

```

```

Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 728 CCTAAGGTCCTCCAGTGTGAATATCAGCGTCTAAGTCAGAAAGGTGCGAGGATGCT 787

Qy 221 TyrProArgGlnIleAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 788 TACCCGAGACAGATAGATGACACATGTTCTGGCGCGGTGACAAAGCAGGTAGAGACTCC 847

Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db 848 TGCAGGGTGATTTCTGGGGGGCCCTGTGGTCTGCAATGGCTCCCTGCGAGGACTGCTGCC 907

Qy 261 TriPLeuAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 908 TGGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGGTCTGTACACGACCTCTGCAAG 967

Qy 281 PheThrLysTriPLeuGlnGluThrIleGlnAlaAsnSer 293
Db 968 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1006

RESULT 4
LOCUS      BD107879      1381 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION 36 human secreted proteins.
ACCESSION  BD107879
VERSION    BD107879.1 GI:23202697
KEYWORDS   JP 2002500035-A/50.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1381)
AUTHORS    Ruben, S.M., Soppet, D.R., Ebner, R., Lafleur, D.W., Ni, J.,
            Brewer, L.A., Olsen, H.S., Duan, R.D. and Rosen, C.A.
TITLE      36 human secreted proteins
JOURNAL    Patent: JP 2002500035-A 50 08-JAN-2002;
            HUMAN GENOME SCIENCES INC
COMMENT     OS Homo sapiens (human)
            PN JP 2002500035-A/50
            PD 08-JAN-2002
            PF 06-JAN-1999 JP 2000527554
            PR 07-JAN-1998 US 60/070657, 07-JAN-1998 US 60/070658 PR
            07-JAN-1998 US 60/070692, 07-JAN-1998 US 60/070704 PI STEVEN
            NI RUBEN, DAVID R SOPPET, REINHARD EBNER, DAVID W LAFLEUR, PI JIAN
            PI LAURIE A BREWER, HENRIK S OLSEN, ROSANNE D DUAN, CRAIG A ROSEN PC
            C12N15/09, A61K31/711, A61K38/00, A61K39/395, A61K39/395, A61K48/00,
            PC A61P5/00, A61P11/06, A61P19/02, A61P29/00, A61P31/18, A61P35/00, PC
            A61P35/02,
            PC A61P37/00, C07K14/435, C07K16/18, C12N1/15, C12N1/21, C12N5/10, PC
            C12N15/00,
            PC A61K37/02, C12N5/00
            CC 36 human secreted proteins
            FH key Location/Qualifiers
            FT source 1. .1381
            /organism="Homo sapiens (human)"
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"

BASE COUNT 329 a 425 c 353 g 274 t
ORIGIN

Alignment Scores:
Pred. No.: 1.31e-120 Length: 1381
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```

US-09-936-271B-14 (1-293) x BD107879 (1-1381)	
QY	1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
Db	129 ATGGCTACAGCAAGACCCCGTGGATGTGGGTGCTCTGTGCTGATCACAGCCTTGCTT 188
QY	21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db	189 CTGGGGTACAGAGAGATGTTCTGCCCAACATGATGTTCTCTGACCACTCTAAC 248
QY	41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60
Db	249 ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCCGGGGAAGACGCCGGTGC 308
QY	61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db	309 GATGACAGCAGAGCCGCATCATCAATGATCGGATCGGATGATGACACCCAGCGGTG 368
QY	81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db	369 CAGGCGCGCTGTGTAGGCCCAACAGCTACTGCGGGCGGTGTGGTGCATCCA 428
QY	101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
Db	429 CAGTGGCTGTCACGGCCGCCACTGCAGGAAGAAAGTTTCAGAGTCGCTCTCGGCCAC 488
QY	121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db	489 TACTCCTGTCACCACTTATGAATCTGGGACGACATGTTCCAGGGGTCAATCCATC 548
QY	141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
Db	549 CCCCACCTGGCTACTCCACCTGGCCACTTAACGACCTCATGCTCATCAACTGAAC 608
QY	161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db	609 AGAAGAATTCGTCCTCCACTAAAGATGTCAGACCCATCAAGCTCTCTCATGTCTCCT 668
QY	181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
Db	669 GCTGGGCAAGAGTCTGTGTGCTGTGGTGGGGACCAACCAAGAGCCCCCAAGTGCATTC 728
QY	201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db	729 CCTAAGTCTCTCCAGTGCTGAATATCAGGTGCTAAGTCAGAAAAGGTGCGAGGATGCT 788
QY	221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db	789 TACCCGACAGATAGATACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCC 848
QY	241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db	849 TGGCAGGTGATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 908
QY	261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db	909 TGGGGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACAGAACCTCTGCAAG 968
QY	281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db	969 TTCACCAAGTGTATCCAGGAACCACTCCAGGCCAATCC 1007
RESULT 5	
LOCUS	AF168768
DEFINITION	Homo sapiens stratum corneum tryptic enzyme (SCTE) mRNA, complete cds.
ACCESSION	AF168768
VERSION	AF168768.2
KEYWORDS	GI:20153423
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
1 (bases 1 to 1387)	
Brattsand,M. and Egelrud,T.	
Purification, molecular cloning, and expression of a human stratum corneum trypsin-like serine protease with possible function in desquamation	
J. Biol. Chem. 274 (42), 30033-30040 (1999)	
99445563	
10514489	
2 (bases 1 to 1387)	
Brattsand,M. and Egelrud,T.	
Submitted (12-JUL-1999) Public Health and Clinical Medicine, Dermatology and Venereology, Umea University, University Hospital, Umea SE-901 85, Sweden	
3 (bases 1 to 1387)	
Brattsand,M. and Egelrud,T.	
Submitted (16-APR-2002) Public Health and Clinical Medicine, Dermatology and Venereology, Umea University, University Hospital, Umea SE-901 85, Sweden	
Sequence update by submitter	
On Apr 16, 2002 this sequence version replaced gi:6063032.	
FEATURES	
source	
1..1387	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/tissue_type="stratum corneum"	
1..1387	
/gene="SCTE"	
170..1051	
/gene="SCTE"	
/note="SCTE; trypsin-like serine protease; possibly involved in desquamation"	
/codon_start=1	
/product="stratum corneum tryptic enzyme"	
/protein_id="AAF03101.1"	
/db_xref="GI:6063033"	
/translation="MATAPPMMVLCALITALLLVTEHVLANNVSDHPNTVPVS GSNQIDGAGEDARDSSSRITNGSDMHTQPWQAALLRPNGLYCGAVLVHPQW LITAAHCRKKVFRVSLHSHYSLSPVYESGQMGQVKSIPHPGYSHFGHNDMLIKLN RIRPTKDVRIPIVNSHSCPSAGTKLVSGWGTTPQVHPFPKVLQCLNTSVLSOKRCE DAYPRDIDTMECAGDKGRDSCQSGSGPVVCGSLQGLVSGDYPGARPNRPGYT NLCKTKWIQETIQANS"	
BASE COUNT 290 a 437 c 351 g 309 t	
ORIGIN	
Alignment Scores:	
Pred. No.: 1..31e-120 Length: 1387	
Score: 1608.00 Matches: 293	
Percent Similarity: 100.00% Conservative: 0	
Best Local Similarity: 100.00% Mismatches: 0	
Query Match: 100.00% Indels: 0	
DB: 9 Gaps: 0	
US-09-936-271B-14 (1-293) x AF168768 (1-1387)	
QY	1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
Db	170 ATGGCTACAGCAAGACCCCGTGGATGTGGGTGCTCTGTGCTGATCACAGCCTTGCTT 229
QY	21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db	230 CTGGGGTACAGAGAGATGTTCTGCCCAACATGATGTTCTCTGACCACTCTAAC 289
QY	41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60
Db	290 ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCCGGGGAAGACGCCGGTGC 349
QY	61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80



US-09-936-271b-14 (1-293) x BD005362 (1-1499)	
QY	1 MetAlaThrAlaArgProThrPheTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
DB	291 ATGGCTACAGCAAGACCCCTGGATGGTGGCTCTGTGCTCTGTATCACACCTTGCCT 350
QY	21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
DB	351 CTGGGGGTACAGACGATGTTCTCGCCCAACATGATGTTCTCTGTGACCACTCTTAAC 410
QY	41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
DB	411 ACCGTGCTCTGGGAGCAACACGAGCTGGGAGCTGGGCGCGGGAGAACCCCGGTGC 470
QY	61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTTP 80
DB	471 GATGACAGCAGCGCGCATCATCAATGATCCGACTGCGATATGACACACCGCGGTGG 530
QY	81 GlnAlaAlaLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
DB	531 CAGGCGCGCTCTTCTTAAGGCCCAACACGACTCTACTGCGGGCGGTGTGTGTCATCCA 590
QY	101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
DB	591 CAGTGGCTGCTCACGGCGCCCTGCAGGAAAGATTTTCAGAGTCCGCTCGGCCAC 650
QY	121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
DB	651 TACTCCCTGTCCACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGTCAATCAATC 710
QY	141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAsn 160
DB	711 CCCACCTCTGGCTACTCCACCTCGCCACTCTAACGACCTCATGCTCAAACTGAAC 770
QY	161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
DB	771 AGAAGAATTCGTCCCTCAAGATGTCAGACCCATCAACGTCTCTCTCTCTCTCTCT 830
QY	181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
DB	831 GCTGGGACAAAGT 890
QY	201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
DB	891 CCTAAGTCTCTCCAGTGTGAATATCAGCGTGTAAAGTCAGAAAAAGTGCAGAGTGTCT 950
QY	221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
DB	951 TACCCGAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCC 1010
QY	241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB	1011 TGCCAGGTGATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1070
QY	261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
DB	1071 TGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTCTCCAAG 1130
QY	281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
DB	1131 TTCACCAAGTGGATCCAGGAAACCATCCAGGCAACTCC 1169
RESULT 8	
LOCUS	BD107865 1516 bp DNA linear PAT 18-SEP-2002
DEFINITION	36 human secreted proteins.
ACCESSION	BD107865
VERSION	BD107865.1 GI:23202683
KEYWORDS	JP 2002500035-A/36.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	

QY	161 ArgAlaGlyLeuArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
DB	676 AGAAGANTCGTCCCACTAAGATGTCAGACCCATCAACGCTCTCTCAITGTCCCTCT 735
QY	181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
DB	736 GCTGGGACAAAGT 795
QY	201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
DB	796 CTAAGGTCTCTCCAGTGTGAATATCAGCGTGTAAAGTCAGAAAAGTGGCAGATGCT 855
QY	221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
DB	856 TACCCGAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGACACTCC 915
QY	241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB	916 TGCCAGGGTGTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCC 975
QY	261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
DB	976 TGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTCTCAAG 1035
QY	281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
DB	1036 TTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC 1074

BD005362	1499 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	BD005362			
DEFINITION	Protease.			
ACCESSION	BD005362			
VERSION	BD005362.1 GI:18633733			
KEYWORDS	JP 2001501837-A/1.			
SOURCE	unidentified			
ORGANISM	unclassified.			
REFERENCE	1 (bases 1 to 1499)			
AUTHORS	Kitado,H., Yoshikawa,A. and Zaiki,T.			
TITLE	Protease			
JOURNAL	Patent: JP 2001501837-A 1 13-FEB-2001;			
COMMENT	THE PROCTER & GAMBLE CO			
	OS Human keratinocyte			
	PN JP 2001501837-A/1			
	PD 13-FEB-2001			
	PF 03-OCT-1997 JP 199521662			
	PI HARUO KITADO, AKIKAZU YOSHIKAWA, TOMOKO ZAIKI			
	PC C12N15/57, C12N9/64, A61K38/48, A61K7/48, C11D3/386, C07K16/40 CC			

PH	Key	Location/Qualifiers
FT	CDS	(291)..(1172)
FT	mat	(489)..(1172)
FT	5' UTR	(1)..(290)
FT	3' UTR	(1173)..(1499).

FEATURES	Location/Qualifiers
source	i..1499
	/organism="unidentified"
	/mol_type="genomic DNA"
	/db_xref="taxon:32644"
BASE COUNT	310 a 469 c 398 g 322 t
ORIGIN	

Alignment Scores:	
Pred. No.:	1.43e-120
Score:	1608.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
Length:	1499
Matches:	293
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

```

REFERENCE 1 (bases 1 to 1516)
AUTHORS Ruben,S.M., Soppet,D.R., Ebner,R., Lafleur,D.W., NI,J.,
        Brewer,L.A., Olsen,H.S., Duan,R.D. and Rosen,C.A.
TITLE 36 human secreted proteins
JOURNAL Patent: JP 2002500035-A 36 08-JAN-2002;
        HUMAN GENOME SCIENCES INC
COMMENT OS Homo sapiens (human)
        PN JP 2002500035-A/36
        PD 08-JAN-2002
        PF 06-JAN-1999 JP 2000527554
        PR 07-JAN-1998 US 60/070657,07-JAN-1998 US 60/070658 PR
        M RUBEN,DANIEL R Soppet,REINHARD EBNER,DAVID W LAFLEUR, PI JIAN
        NI, LAURIE A BREWER,HENRIK S OLSEN,ROSANNE D DUAN,CRAIG A ROSEN PC
        C12N15/09,A61K31/711,A61K38/00,A61K39/395,A61K48/00,
        PC A61P5/00,
        PC A61B7/00,A61P11/06,A61P19/02,A61P29/00,A61P31/18,A61P35/00, PC
        A61P35/02,
        PC A61P37/00,C07K14/435,C07K16/18,C12N1/15,C12N1/21,C12N5/10, PC
        C12N15/00,
        PC A61K37/02,C12N5/00
        CC n equals a,t,g, or c
        FH Key Location/Qualifiers
        FT source 1..1516
        FT /organism='Homo sapiens (human)'.
FEATURES
source
    1..1516
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
BASE COUNT 338 a 472 c 381 g 324 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.45e-120 Length: 1516
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-936-271b-14 (1-293) x BD107865 (1-1516)
Qy 1 MetaLthrAlaArgProTropMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
Db 254 ATGGCTACAGCAAGACCCCTGGTGTGGTGTCTGTCTGTCTGTATCAGACCTTCTT 313
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 314 CTGGGGGTACAGAGCATGTTCTCGCCCAACAATGATGTTCTCTGTGACCCCTCTAAC 373
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 374 ACCGTGCCCTTGGAGCAACAGACCTGGAGCTGGGCGGGGGAAGAGCGCCGGTGC 433
Qy 61 AspAspSerSerArgIleAlaAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 434 GATCAGCAGCAGCAGCCGCATCATCAATGATCCGACTCGCATATGACACCCAGCCGTGG 493
Qy 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 494 CAGCCCGCGTGTGGTAAAGCCCAACAGCTCTACTCTACTCTGGGGGGGGTGTGGTGCATCCA 553
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
Db 554 CAGTGGGTGTCACGGCGGCCACTGCAGGAAGAAAGTTTTCAGAGTCCGTCTCGGCCAC 613
Qy 121 TyrSerLeuSerProValTyrGluSerGlnGlnMetPheGlnGlyValLysSerIle 140
Db 614 TACTCCCTGTCCACAGTTTATGATCTGGCAGCAGATGTTCCAGGGGGTCAATCCATC 673
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160

```

---

```

674 CCCACCCCTGGTACTCCACCCCTGGCCACTTAACGACCTCATGCTCAAACTGAAC 733
161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
734 AGAAGAATTCGTCCCACTAAAGATGTCAGACCCCACTCAACGCTCTCTCTCATTTGTCCTCT 793
181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200
794 GCTGGGACAAAGTCTTGGTGTCTGGCTGGGGGACAAACCAAGAGCCGCCCAAGTGCACCTC 853
201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
854 CCTAAGGTCTCTCCAGTCTTGAATATATCAGGCTGTAAGTCAGAAAAGGTCGAGGATGCT 913
221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
914 TACCCGAGACAGATAGATGACACCATGTTCTGCCCGCGGTGACAAAGCAGGTAGACACTCC 973
241 CysGlnGlyAspSerGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
974 TGCAGGGTGTATCTGGGGGCGCTGTGTCTGCAATGGCTCCCTGCAGGACTCGTCTCC 1033
261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
1034 TGGGGAGATTACCTTGTGCGGCCCAACAGACCGGGTGTCTTACAGAACTCTGCAAG 1093
281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
1094 TTCACCAAGTGGATCCAGGAACCCNTCCAGGCCAANTCC 1132
RESULT 9
BC008036 1536 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, kallikrein 5, clone MGC:1107 IMAGE:2989806, mRNA,
DEFINITION complete cds.
ACCESSION BC008036
VERSION BC008036.1 GI:14165477
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1536)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettman and Anuradha Madan
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 3 Row: m Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6912643.
Location/Qualifiers
1..1536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="locusID:25818"
/db_xref="taxon:9606"
FEATURES
source
    1..1536

```

/clone="MGC:1107 IMAGE:2989806"  
 /tissue\_type="Ovary, adenocarcinoma"  
 /clone\_lib="NIH\_MGC\_9"  
 /lab\_host="DH10B-R"

/note="Vector: pOTB7"

CDS

294..1175

/codon\_start=1

/product="kallikrein 5"

/protein\_id="AAH08036.1"

/db\_xref="GI:14165478"

/translation="MATAAPPMWVLCALITALLLGVTREHVLANNVDSCDHPSTNTPVS  
 GSNQDAGAGDEADSDSSRIINGSDCMHTQWQAALLLRPNOLYCGAVLVHPQW  
 LITAAHCRKRVRLGHYSLSPVYESGQMFQYKSIPIHPGYSPHGHNDLMLIKLN  
 RIRPTKDRPINVSHCPKSAGTKCLVSGWGTTPQVHFPRKVLQCLNLSVLSQRCE  
 DAYPRIDDTMFCAGDKAGRDSCQGDGSGPVVNGSLQGLVSGWDYPCARPNRPGVYT  
 NLCKETKWLQETIQANS"

BASE COUNT 346 a 467 c 401 g 322 t

ORIGIN

Alignment Scores:

Pred. No.: 1,47e-120 Length: 1536  
 Score: 1608.00 Matches: 293  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-936-271B-14 (1-293) x BC008036 (1-1536)

QY 1 MetAlaThraAlaArgProProTrpMetTrpValLeuCysAlaLeuLeuThrAlaLeuLeu 20  
 Db 294 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTGCTGATCAGACCTTGCTT 353  
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
 Db 354 CTGGGGGTACAGAGCATGTTCTGCCAACAATGATGTTCTGTGACCACTCTTAAC 413  
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
 Db 414 ACCGTGCCCTCTGGAGAACACAGGACCTGGGAGCTGGGGCCGGGAAACGCCGTG 473  
 QY 61 AspAspSerSerArgIleLeuAsnGlySerAspCysAspMethHisThrGlnProTrp 80  
 Db 474 GATGACAGCAGCAGCGCATCAATGATGATGATGATGATGATGATGATGATGATG 533  
 QY 81 GlnAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
 Db 534 CAGGCCGCGCTGTGTAAAGGCCCAACACGCTCTACTGCGGGCGGTGTGTGTCATCCA 593  
 QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120  
 Db 594 CAGTGGCTGCTCAGCGCCGCCACTCGCAGGAGAAAGTTTTCAGAGTCCGTCTCGGCC 653  
 QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
 Db 654 TACTCCCTGTCCACAGTTTATGAATCTGGGCAGCAGATGTTCCAGGGGTCAATCCATC 713  
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160  
 Db 714 CCCACCCCTGGTACTCCACCCCTGGCCACTCTAAGCAGCTCATGCTCATCAACCTGAAC 773  
 QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
 Db 774 AGAAGAAATTCGCCCACTAAAGATGTCAGACCCATCAAGCTCTCTCTCATGTCCCTCT 833  
 QY 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
 Db 834 GCTGGGACAAAGTCTGTGTGCTGGTGGGGACACCAAGAGCCCTCAAGTGCATTC 893  
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
 Db 894 CCTAAGTCTCTCCAGTGCTTGAATATCAGCGTGCTAAGTCAGAAAAAGGTGCGAGGATGCT 953

QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
 Db 954 TACCCGAGACAGATAGATCACCATTGTTCTGGCGGGTGACAAACAGGTAGACTCC 1013  
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
 Db 1014 TGCCAGGGTGATTTCTGGGGGCTGTGTGCTGCAATGGCTCCCTGAGGAGCTCGTGTCC 1073  
 QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
 Db 1074 TGGGGAGATTACCTTGTGCGCGCCCAACACCGGGTGTCTACAGAACCTCTCAAG 1133  
 QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
 Db 1134 TTCACCAAGTGGATCCAGGAACCATCCAGGCCAATCC 1172  
 RESULT 10  
 AR252580  
 LOCUS AR252580 1570 bp DNA linear PAT 20-DEC-2002  
 DEFINITION Sequence 308 from patent US 6478825.  
 ACCESSION AR252580  
 VERSION AR252580.1 GI:27300488  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1570)  
 AUTHORS Winterbottom,J.M., Shimp,L., Boyce,T.M. and Kaes,D.  
 TITLE Implant, method of making same and use of the implant for the  
 treatment of bone defects  
 JOURNAL Patent: US 6478825-A 308 12-NOV-2002;  
 FEATURES Location/Qualifiers  
 source 1..1570  
 BASE COUNT 332 a 485 c 424 g 329 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,5e-120 Length: 1570  
 Score: 1608.00 Matches: 293  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-936-271B-14 (1-293) x AR252580 (1-1570)  
 QY 1 MetAlaThraAlaArgProProTrpMetTrpValLeuCysAlaLeuLeuThrAlaLeuLeu 20  
 Db 354 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTGCTGATCAGACCTTGCTT 413  
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
 Db 414 CTGGGGGTACAGAGCATGTTCTGCCAACAATGATGTTCTGTGACCACTCTTAAC 473  
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
 Db 474 ACCGTGCCCTCTGGGAGAACACAGGACCTGGGAGCTGGGGCCGGGAAACGCCGTG 533  
 QY 61 AspAspSerSerArgIleLeuAsnGlySerAspCysAspMethHisThrGlnProTrp 80  
 Db 534 GATGACAGCAGCAGCGCATCAATGATGATGATGATGATGATGATGATGATGATG 593  
 QY 81 GlnAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
 Db 594 CAGGCCGCGCTGTGTAAAGGCCCAACACGCTCTACTGCGGGCGGTGTGTGTCATCCA 653  
 QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120  
 Db 654 CAGTGGCTGCTCAGCGCCGCCACTCGCAGGAGAAAGTTTTCAGAGTCCGTCTCGGCC 713  
 QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140

```

Db      714  TACTCCCTGTCACACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGGTCAAAATCCATC 773
Qy      141  ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuHisLeuAsn 160
Db      774  CCCACCCCTGGTACTCCACCTCCGCACTCTTAACGACCTCATGTCATCAACATGAAC 833
Qy      161  ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db      834  AGAAGAAATTCGTCACCAATAAGATGTAGACACCATCAACGCTCTCTCTCTCTCTCTCT 893
Qy      181  AlacGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200
Db      894  GCTGGGACAAGTCTGGTGTCTGGCTGGGGGACAAACCAAGACCCCAAGTGCACCTC 953
Qy      201  ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db      954  CCTAAGGTCTCCAGTCTTGAATATCAGCGTCTAAGTCAGAAAGTTCGAGGATGCT 1013
Qy      221  TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db      1014  TACCCGAGACAGATAGATGACACCATGTTCTGGCCCGGTGACAAAGCAGGTAGAGACTC 1073
Qy      241  CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db      1074  TGCAGGGTGATCTGGGGGGCTTGCTGCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1133
Qy      261  TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db      1134  TGGGAGATTTACCTTGTGCGCGCCCAACAGACCCGGGTGCTACAGGAACCTCTGCAAG 1193

Qy      281  PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db      1194  TTCACCAAGTGGATGCCAGAAACCATCCAGGCCCAACTCC 1232

RESULT 11
LOCUS   AX080829                      1570 bp      DNA      linear      PAT 27-FEB-2001
DEFINITION   Sequence 75 from Patent WO0109327.
ACCESSION   AX080829
VERSION     AX080829.1  GI:13169796
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Ashkenazi,A.J., Baker,K.P., Goddard,A., Godowski,P.J., Gurney,A.L.,
            Kijavini,I.J., Lafleur,M., Mark,M.R., Marsters,S.A., Pitti,R.M.,
            Watanabe,C.K. and Wood,W.I.
TITLE       Method of preventing the injury or death of retinal cells and
            treating ocular diseases
JOURNAL     Patent: WO 0109327-A 75 08-FEB-2001;
            Genentech, Inc. (US)
FEATURES
  source
    1..1570
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
BASE COUNT  332 a 485 c 424 g 329 t
ORIGIN
Alignment Scores:
Pred. No.:      1.5e-120      Length:      1570
Score:          1608.00      Matches:      293
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6           Gaps:      0

US-09-936-271b-14 (1-293) x AX080829 (1-1570)
Qy      1  MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
            |||||

```

```

Db      354  ATGGCTACACAAAGACCCCTCGATGTGGGTGCTCTGTGTGCTCTGTATCAGAGCTTGCTT 413
Qy      21   LeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCysAspHisProSerAsn 40
            |||||
Db      414  CTGGGGTTCACAGAGATGTTCTCGCCAAATGATGTTCTTCTGTGACCACTCTTAAC 473
Qy      41   ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
            |||||
Db      474  ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGGGAAGACGCCGGTCG 533
Qy      61   AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
            |||||
Db      534  GATGACAGCAGCAGCGCATCATCAATGGATCCGACTGCGCATATGACACCCAGCGGTGG 593
Qy      81   GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
            |||||
Db      594  CAGGCCGCGTGTGCTAAGGCCCAACCACTCTACTGCGGGGGTGTGGTGGTCATCCA 653
Qy      101  GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120
            |||||
Db      654  CAGTGGGTGCTCAGCGCGCCACTGCGAGGAAGAAAGTTTTCAGAGTCCGCTCTCGGCCAC 713
Qy      121  TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
            |||||
Db      714  TACTCCTGTCCACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGGTCAAAATCCATC 773
Qy      141  ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160
            |||||
Db      774  CCCACCCCTGGTACTCCCACTCCGCACTCTTAACGACCTCATGCTCATCAAACTGAAC 833
Qy      161  ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
            |||||
Db      834  AGAAGAAATTCGTCGCCACTAAAGATGTGAGACCCATCAACGCTCTCTCTCTCTCTCTCT 893
Qy      181  AlacGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200
            |||||
Db      894  CTTGGGACAAGTCTGGTGTCTGGCTGGGGGACAAACCAAGACCCCAAGTGCACCTTC 953
Qy      201  ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
            |||||
Db      954  CCTAAGGTCTCCAGTCTTGAATATCAGCGTGTAAAGTTCAGAAAGTTCGAGGATGCT 1013
Qy      221  TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
            |||||
Db      1014  TACCCGAGACAGATAGATGACACCATGTTCTGGCCCGGTGACAAAGCAGGTAGAGACTC 1073
Qy      241  CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
            |||||
Db      1074  TGCAGGGTGATCTGGGGGGCTTGCTGCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1133
Qy      261  TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
            |||||
Db      1134  TGGGGAGATTTACCTTGTGCGCGCCCAACAGACCCGGGTGCTACAGGAACCTCTGCAAG 1193
Qy      281  PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
            |||||
Db      1194  TTCACCAAGTGGATGCCAGAAACCATCCAGGCCCAACTCC 1232

RESULT 12
AX03421
LOCUS   AX03421                      1570 bp      DNA      linear      PAT 14-JUN-2002
DEFINITION   Sequence 308 from Patent WO0073454.
ACCESSION   AX03421
VERSION     AX03421.1  GI:21436934
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Ashkenazi,A.J., Baker,K.P., Botstein,D., Desnoyers,L., Eaton,D.,
            Ferrara,N., Gerber,H., Gerritsen,M., Goddard,A., Godowski,P.,
            Grimaldi,C.J., Gurney,A.L., Kijavini,I., Napier,M.A., Pan,J.,
            ...

```



Paoni, N.F., Roy, M., Stewart, T.A., Tumas, D., Watanabe, C.K.,  
Williams, P., Wood, W.I. and Zhang, Z.  
Secreted and transmembrane polypeptides and nucleic acids encoding  
the same

JOURNAL Patent: WO 0073454-A 308 07-DEC-2000;

Genentech Inc. (US)

## FEATURES

Location/Qualifiers

1..1570

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 332 a 485 c 424 g 329 t

## ORIGIN

Alignment Scores:  
Pred. No.: 1..1570 Length: 1570  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-936-271B-14 (1-293) x AX403421 (1-1570)

QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
DB 354 ATGGCTACAGCAAGACCCCTGGATGGTGCTCTGTCTGATCAGCGCTTGCTT 413  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
DB 414 CTGGGGGTACAGAGCATGTTCTCGCAACAATGATGTTCTGTGACCACTCTTAAC 473  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60  
DB 474 ACCGTGCCCTCTGGGAGCAACAGACCTGGGAGCTGGGCGGGGAAGCGCGGTGC 533  
QY 61 AspAspSerSerArgIleLeuAsnGlySerAspCysAspMethIleThrGlnProTrp 80  
DB 534 GATGACAGCAGCAGCGCATCATCAATGGATCGGACTGGCATATGCACACCCGCGTG 593  
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCysGlyAlaValLeuValHisPro 100  
DB 594 CAGCGCGCGCTGTGTAAAGCCCAACAGCTCTACTGCGGGCGGTGTGGTGATCCA 653  
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120  
DB 654 CAGTGGCTGCTCAGCGCGCCCACTGCAGGAAGAAAGTTTCAGAGTCCGCTCGGCCAC 713  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
DB 714 TACTCCCTGTCCAGCTTTATGATCTGGGCACAGATGTTCCAGGGGTCAATCCATC 773  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160  
DB 774 CCCACCCCTGGCTACCCACCCCTGGCCACTCTAAGACCTCATGCTCATCAACTGAAC 833  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
DB 834 AGAAGAAATGCTCCCACTAAGATGTCAGCCCATCAACGCTCCCTCATTTGCTCCCT 893  
QY 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200  
DB 894 GCTGGGCAAAAGTGTCTGGTCTGGCTGGGGGACAAACCAAGAGCCCCCAAGTGCACTTC 953  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
DB 954 CCTAAGTCTCTCCAGTCTTGAATATCAGCGTCTTAAGTCAGAAAAAGGTGCGAGGATGCT 1013  
QY 221 TyrProArgGlnIleAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
DB 1014 TACCCGAGACAGATAGATACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGATCC 1073  
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260

DB 1074 TGCAGGGTGATTTCTGGGGCGCTGTGGTCTGCAATGGCTCCCTCAGGAGCTGCTGTC 1133  
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValThrAsnLeuCysLys 280  
DB 1134 TGGGGAGATTACCTTTGTGCCCGCCCAACAGACCGGGTGTCTACAGAACCTCTCAAG 1193  
QY 281 PheThrLysTrpIleGlnGlnThrIleGlnAlaAsnSer 293  
DB 1194 TTCACCAAGTGGATCCAGGNAACCATCCAGGCCAATCC 1232

## RESULT 13

AX464322

LOCUS

DEFINITION

Sequence 455 from Patent WO0140466.

AX464322

VERSION

AX464322.1 GI:21899177

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

Baker, K.P., Beresini, M., Deforge, L., Desnoyers, L., Filvaroff, E.,

Gao, W.Q., Gerritsen, M.E., Goddard, A., Godowski, P.J., Gurney, A.L.,

Sherwood, S., Smith, V., Stewart, T.A., Tumas, D., Watanabe, C.K.,

Wood, W.L. and Zhang, Z.

SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING

THE SAME

JOURNAL

Patent: WO 0140466-A 455 07-JUN-2001;

Genentech Inc. (US)

FEATURES

Location/Qualifiers

1..1570

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 332 a 485 c 424 g 329 t

## ORIGIN

## Alignment Scores:

Pred. No.: 1..1570 Length: 1570  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-936-271B-14 (1-293) x AX464322 (1-1570)

QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
DB 354 ATGGCTACAGCAAGACCCCTGGATGGTGCTCTGTCTGATCAGCGCTTGCTT 413  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
DB 414 CTGGGGGTACAGAGCATGTTCTCGCAACAATGATGTTCTGTGACCACTCTTAAC 473  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60  
DB 474 ACCGTGCCCTCTGGGAGCAACAGACCTGGGAGCTGGGCGGGGAAGCGCGGTGC 533  
QY 61 AspAspSerSerArgIleLeuAsnGlySerAspCysAspMethIleThrGlnProTrp 80  
DB 534 GATGACAGCAGCAGCGCATCATCAATGGATCCGACTGCGATATGCACACCCGCGTG 593  
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCysGlyAlaValLeuValHisPro 100  
DB 594 CAGCGCGCGCTGTGTAAAGCCCAACAGCTCTACTGCGGGCGGTGTGGTGATCCA 653  
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120  
DB 654 CAGTGGCTGCTCAGCGCGCCCACTGCAGGAAGAAAGTTTCAGAGTCCGCTCGGCCAC 713

```
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 714 TACTCCCTGTCCACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGTCAAAATCCATC 773
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
Db 774 CCCACACCTGGCTACTCCCACTGGCCACTCTAAGCAGCATCAAGCTCATGCTCAAACTGAAC 833
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 834 AGAAGATTCTGCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCTCATTTGCCCTCT 893
Qy 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
Db 894 GCTGGGACAAAGTCTGGTGTCTGGCTGGGGGACAAACCAAGACCCCAAGTGCACCTTC 953
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 954 CCTAAGGTCTCCAGTCTTGAATATCAGGCTGCTAAGTCAGAAAGGTGCGAGGATGCT 1013
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 1014 TACCCGAGACAGATAGATGACACCATCTTCTGCGCGGTGACAAAGCAGGTAGAGACTCC 1073
Qy 241 CysGlnGlyaspSerGlyGlyProValValCysasnGlySerLeuGlnGlyValSer 260
Db 1074 TGCCAGGGTGATTTCTGGGGGCGCTGTGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1133
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 1134 TGGGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTGCTACAGAACCTCTGCAAG 1193
Qy 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 1194 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1232
RESULT 14
LOCUS AR078184 1476 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 2 from patent US 5962300.
ACCESSION AR078184
VERSION AR078184.1 GI:10004930
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Hillman,J.L. and Lal,P.
TITLE Human kallikrein
JOURNAL Patent: US 5962300-A 2 05-OCT-1999;
FEATURES Location/Qualifiers
source 1..1476
/organism="unknown"
BASE COUNT 303 a 466 c 388 g 319 t
ORIGIN
Alignment Scores:
Pred. No.: 1.57e-119 Length: 1476
Score: 1595.00 Matches: 293
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 0
Query Match: 99.19% Indels: 1
DB: 6 Gaps: 0
US-09-936-271B-14 (1-293) x AR078184 (1-1476)
Qy 1 MetalathrAlaArgProProThrMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
Db 277 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTCTGTGATCAGACGCTTGTCT 336
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 337 CTGGGGGTACAGAGATGTTCTCGCCCAACAATGATGTTCTCTGTGACCACCCCTCTAAC 396
```

```
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 397 ACCGTGCCCTCTGGGAGCAACACGAGCCTGGGAGCTGGGGCGGGGAAGACGCCCGGTG 456
Qy 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 457 GATGACAGCAGCAGCGCATCATATGGATCGACTGCGATATGACACCCAGCGGTGG 516
Qy 81 -GlnAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPr 100
Db 517 CGAGGCGCGCTGTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTTGGTCATCC 576
Qy 100 oGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyH 120
Db 577 ACAGTGGCTGCTCACGGCGCCCACTGCAGGAAGAAAGTTTTCAGAGTGGCTCTCGCCA 636
Qy 120 sTyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerI 140
Db 637 CTACTCCCTGTCCAGTCTTATGAATCTGGCAGCAGATGTTCCAGGGGTCAAAATCCAT 696
Qy 140 eProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAs 160
Db 697 CCCCACCTGGCTACTCCCACTGCGCACTCTAACGACCTCATGCTCATCAAACTGAA 756
Qy 160 nArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSe 180
Db 757 CAGAAGAATTCGTCCCACTAAAGATGTCAGACCCATCAAGCTCTCTCTCATTTGTCCTC 816
Qy 180 rAlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPh 200
Db 817 TGCTGGGACAAAGTCTGTGTGTCGCGTGGGGGACCAACCAAGAGCCCCCAAGTGCATT 876
Qy 200 eProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAl 220
Db 877 CCCTAAGGTCTCCAGTCTTGAATATCAGGCTGCTAAGTCAGAAAAAGTTCGAGAGTGC 936
Qy 220 aTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSe 240
Db 937 TTACCCGAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTC 996
Qy 240 rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSe 260
Db 997 CTGCCAGGTGATTCCTGGGGGCGCTGTGTGCTGCAATGGCTCCCTGCGAGGACTCGTGC 1056
Qy 260 rTTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLy 280
Db 1057 CTGGGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTCTCTACAGAACCTCTCGAA 1116
Qy 280 sPheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 1117 GTTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1156
RESULT 15
LOCUS AR137506 1476 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 2 from patent US 6197511.
ACCESSION AR137506
VERSION AR137506.1 GI:14479015
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Hillman,J.L. and Lal,P.
TITLE Human kallikrein
JOURNAL Patent: US 6197511-A 2 06-MAR-2001;
FEATURES Location/Qualifiers
source 1..1476
/organism="unknown"
BASE COUNT 303 a 466 c 388 g 319 t
ORIGIN
```

## Alignment Scores:

Pred. No.:	1.57e-119	Length:	1476
Score:	1595.00	Matches:	293
Percent Similarity:	99.66%	Conservative:	0
Best Local Similarity:	99.66%	Mismatches:	0
Query Match:	99.19%	Indels:	1
DB:	6	Gaps:	0

US-09-936-271B-14 (1-293) x AR137506 (1-1476)

QY	1	MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
DB	277	ATGGCTACAGCAGACAGCCCTGGATGTGGGTCTCTGTGCTGATCACAGCCTTGCTT 336
QY	21	LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
DB	337	CTGGGGGTACAGAGCATGTTCTGCCAACACATGATGTTCTGTGACCACTCCCTCTAAC 396
QY	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
DB	397	ACCGTGCCTCTGGGAGCAACAGGACCTGGGAGCTGGGCGGGGGAAGACGCCCGGTGCG 456
QY	61	AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMethIsthrGlnProTrp 80
DB	457	GATGACAGCAGCAGCCGCATCATCAATGGATCCGACTGGCATATGCACACCCAGCCGTGG 516
QY	81	-GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPr 100
DB	517	GCAGGCGGGGTGCTTAAAGGCCCAACACAGCTCTACTCGGGGGCGGTGTGGTGATCC 576
QY	100	oGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyH1 120
DB	577	ACAGTGGCTGCTCACGCCGCCACATGCAGGAAGAAAGTTTTCAGAGTCCGCTCTCGGCCA 636
QY	120	sTyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerI1 140
DB	637	CTACTCCCTGTCAACCAATTTATGAATCTGGGAGCAGATGTTCCAGGGGGTCAAAATCCAT 696
QY	140	eProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAs 160
DB	697	CCCCACCTGGGTACTCCACCCTGGCCACTCTAACGACCTCATGCTCATCAACTGAA 756
QY	160	nArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSe 180
DB	757	CAGAGAATTCGTCACCTAAAGATGTCAGACCCATCAACGCTCTCTCTCAITGTCCCTC 816
QY	180	rAlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPh 200
DB	817	TGCTGGGCAAAAGTGTGGTGTCTGGGTGGGGGACAAACAGAGCCCCCAAGTGCACCT 876
QY	200	eProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAl 220
DB	877	CCCTAAGGTCTCCAGTGTGTAATATCAGGTGCTAAGTCAGAAAAGGTGCAGGATGC 936
QY	220	aTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSe 240
DB	937	TTACCCGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGACACTC 996
QY	240	rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSe 260
DB	997	CTGCCAGGTGATTTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTG 1056
QY	260	rTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLy 280
DB	1057	CTGGGGAGATATACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAA 1116
QY	280	sPheThrLysTrpIleGlnGlnThrIleGlnAlaAsnSer 293
DB	1117	GTTCAACAAGTGGATCCAGGAAACCATCCAGGCCCACTCC 1156

Search completed: August 5, 2003, 13:24:21  
Job time : 3136 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 5, 2003, 05:55:25 ; Search time 274 Seconds  
(without alignments)  
2886.625 Million cell updates/sec

Title: US-09-936-271b-14  
Perfect score: 1608  
Sequence: 1 MATARPPMVMVLCALITALL.....VYTNLCKTKTKWQETIQANS 293

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues  
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US09936271/runat\_04082003\_104245\_25695/app\_query.fasta\_1.455  
-DB=N\_Geneseq\_19Jun03 -QFMT=fastp -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0  
-LIST=45 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-MODE=LOCAL -OUTPMT=pto -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-USPR=US09936271 -CGN\_1\_1\_246 -runat\_04082003\_104245\_25695 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*  
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1608	100.0	882	20	AAZ23318	Human PDSP-1 DNA c
2	1608	100.0	1302	20	AAZ57989	Human BS247 specif
3	1608	100.0	1302	20	AAZ57990	Human BS247 specif
4	1608	100.0	1381	20	AAZ06259	Human secreted pro
5	1608	100.0	1499	20	AAZ60578	Human keratinocyte
6	1608	100.0	1516	20	AAZ06245	Human secreted pro
7	1608	100.0	1539	20	AAZ23317	Human PDSP-1 DNA.
8	1608	100.0	1570	21	AAZ65070	Membrane-bound pro
9	1608	100.0	1570	22	AAZ21471	Human cDNA sequenc
10	1608	100.0	1570	22	AAZ44216	Human PRO1132 (UNQ
11	1608	100.0	1570	24	ABK28605	Human DNA59767-148
12	1608	100.0	1570	25	ABK03830	cDNA encoding huma
13	1608	100.0	1570	25	ACA04251	Human cDNA encodin
14	1608	100.0	1570	25	ABX89368	DNA encoding novel
15	1608	100.0	1570	25	ABX80318	Human secreted or
16	1608	100.0	1570	25	ABX80822	Human secreted tra
17	1608	100.0	1570	25	ABX81205	Human secreted or
18	1608	100.0	1570	25	ABX90295	Human secreted tra
19	1608	100.0	1570	25	ABX77906	Human PRO polynucl
20	1608	100.0	1570	25	ABX79502	Human secreted/tr
21	1608	100.0	1570	25	ABX64141	cDNA encoding huma
22	1608	100.0	1570	25	ABX17105	Human PRO polynucl
23	1595	99.2	1476	20	AAZ16295	Human kallikrein e
24	1513	94.1	1439	20	AAZ06260	Human secreted pro
25	1403	87.3	1504	24	ABK30231	Human G-protein-co
26	1002	62.3	11570	21	AAZ95905	Human KLK-L2 gene.
27	734	45.6	455	20	AAZ57986	Human BS247 specif
28	725	45.1	735	21	AAZ79469	cDNA sequence of h
29	725	45.1	735	24	ABK29013	Human breast tumou
30	725	45.1	738	20	AAZ84240	DNA encoding human
31	688	42.8	673	22	AAF44901	Human breast cance
32	671	41.7	765	22	AAZ63914	Human prostate cDN
33	671	41.7	765	22	AAH93821	Human prostate-spe
34	671	41.7	765	22	AAH85135	Human prostate-spe
35	671	41.7	765	22	ABU95285	Human P703P putati
36	671	41.7	765	25	ACA59722	Prostate cancer th
37	671	41.7	765	25	ABQ83343	Human KLK4 encodin
38	654.5	40.7	1052	21	AAZ87798	Activation constru
39	654.5	40.7	1052	22	AAZ55270	Nucleotide sequenc
40	652.5	40.6	1146	20	AAZ84589	Human secreted pro
41	652.5	40.6	1146	22	ABA83430	Human secreted pro
42	652.5	40.6	1166	22	AAZ14841	Human PS133 consen
43	650.5	40.5	833	19	AAZ42925	DNA encoding a hum
44	650.5	40.5	1106	20	AAZ22638	CASB12 nucleotide
45	650.5	40.5	1158	20	AAZ22639	CASB12 derived fro

ALIGNMENTS

RESULT 1  
AAZ23318  
ID AAZ23318 standard; DNA; 882 BP.  
AC AAZ23318;  
XX  
XX  
XX 03-DEC-1999 (first entry)  
XX Human PDSP-1 DNA coding region.  
XX  
XX PDSP-1; Tango 114; prostate-derived serine protease; fusion protein;  
KW treatment; proliferative disorder; prostate cancer; screening;  
KW predictive medicine; growth factor; growth factor biosynthesis;  
KW cellular proliferation; growth factor binding protein; metastasis;  
KW cellular differentiation; prostate development; detection; forensic;  
KW serine protease; human; ds.  
XX

OS Homo sapiens.  
 PN WO9946391-A2.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 12-MAR-1999; 99WO-US05416.  
 XX  
 PR 12-MAR-1998; 98US-0041400.  
 XX  
 PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.  
 PI Holtzman DA;  
 XX  
 DR WPI; 1999-561678/47.  
 DR P-PSDB; AAY30524.  
 XX  
 PT Novel prostate derived serine protease polynucleotides and polypeptides  
 PT used to modulate cellular processes  
 XX  
 PS Claim 1a; Fig 1A; 94pp; English.  
 XX  
 CC This invention describes a novel human prostate-derived serine protease  
 CC (PDSP) polynucleotide (also known as Tango 114). The prostate derived  
 CC serine protease (PDSP) polypeptide is used to modulate a variety of  
 CC cellular processes. It can be used to produce fusion proteins. PDSP is used  
 CC to treat proliferative disorders, e.g. prostate cancer. The protein may  
 CC also be used to produce antibodies, and to identify antagonists and  
 CC agonists. The PDSP polynucleotides, polypeptides, homologs and antibodies  
 CC can be used in screening assays; predictive medicine; and methods of  
 CC treatment. PDSP cleaves growth factors, and can be used for the  
 CC modulation of growth factor biosynthesis; generation of active peptides;  
 CC regulation of cellular proliferation; degradation of growth factor  
 CC binding proteins; regulation of cellular differentiation; regulation of  
 CC metastasis; and regulation of prostate development. The PDSP  
 CC polynucleotides can be used to express the protein; to detect PDSP mRNA;  
 CC to detect genetic alterations in the PDSP gene; in forensic biology; and  
 CC as a source of primers and probes. As serine protease have important  
 CC roles in cellular processes, there exists a need for identifying novel  
 CC serine protease such as the prostate derived serine protease (PDSP) of  
 CC the invention. This sequence encodes the human PDSP-1 described in the  
 CC invention.  
 XX  
 SQ Sequence 882 BP; 193 A; 272 C; 237 G; 180 T; 0 other;

Alignment Scores:  
 Pred. No.: 9,77e-138 Length: 882  
 Score: 1608.00 Matches: 293  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-09-936-271B-14 (1-293) x AAZ23318 (1-882)

QY 1 MetaAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
 DB 1 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTCTGTATCATCAGACGCTTGTCT 60  
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
 DB 61 CTGGGGGTTCACAGAGCATGTTCTCGCAACAATGATGTTCTCTGTGACCCCTCTAAC 120  
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60  
 DB 121 ACCGTGCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGGGAAGACGCGGTCG 180  
 QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80  
 DB 181 GATGACAGCAGCAGCCGCATCATCAATGATCCGACTGCGATATGACACCCACCCCTGG 240  
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

241 CAGGCCGCGCTGTTGCTAAGGCCCAACAGCTCTACTCGGGGGCGGTGTTGGTGCATCCA 300  
 QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
 DB 301 CAGTGGCTGCTCAGCGCCGCCCTACTGAGGAAGAAGTTTTTCAGAGTCCGTCTCGGCAC 360  
 QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
 DB 361 TACTCCCTGTCAACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGGTCAAACTCATC 420  
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
 DB 421 CCCACCCCTGGCTACTCCACCCCTGGCAGCTTAAAGATGTGAGATGCTCACTCATCAAACTGAAC 480  
 QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerHisCysProSer 180  
 DB 481 AGAAGAATTCGTCCACTAAAGATGTGAGATGCTCACTCATCACTCATCTCATTTGCCCTCT 540  
 QY 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
 DB 541 GCTGGACAAAGTCTTGGTCTGTGGCTGGGGGACACCAAGAGACCCCAAGTGCACCTTC 600  
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
 DB 601 CCTAAGGCTCTCCAGTCTTGAATATCAGCGTGTAAAGTCAGAAAGGTGCGAGATGCT 660  
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
 DB 661 TACCCGAGACAGATAGATGACCATGTTCTGCGCCCGGTGACAAGCAGGTAGAGACTCC 720  
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
 DB 721 TGCCAGGCTGATTTCTGGGGGCCCTGTGTCTGCAATGGCTCCCTGCAGGACTCGTGTC 780  
 QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
 DB 781 TGGGGAGATTACCCCTGTGCGCCGCCCAACAGACGGGTGTCTACACAACTCTGCAAG 840  
 QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
 DB 841 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 879

RESULT 2  
 AAX57989  
 ID AAX57989 standard; DNA; 1302 BP.  
 XX  
 AC AAX57989;  
 XX  
 DT 19-JUL-1999 (first entry)  
 XX  
 DE Human BS247 specific polynucleotide #7.  
 XX  
 KW BS247; detection; diagnosis; breast cancer; atypical hyperplasia;  
 KW fibroadenoma; cystic breast disease; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9922027-A1.  
 XX  
 PD 06-MAY-1999.  
 XX  
 PF 28-OCT-1998; 98WO-US22906.  
 XX  
 PR 28-OCT-1997; 97US-0968838.  
 PR 28-OCT-1997; 97US-0063431.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;  
 PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;  
 PI Russell JC, Stroupe SD, Yu H;  
 XX  
 DR WPI; 1999-312977/26.



## Alignment Scores:

Pred. No.: 1,62e-137 Length: 1302  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-09-936-271B-14 (1-293) x AA57990 (1-1302)

Qy 1 MetalThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
Db 88 ATGGCTACAGAACACCCCTGGATGGTGGTCTGTCTGTATCAGACCTTGCTT 147  
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
Db 148 CTGGGGTTCACAGACATGTTCTCGCCCAATGATGTTTCTGTGACCACTCTAAC 207  
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60  
Db 208 ACCGTGCTCTGGGAGCAACAGGACCTGGAGCTGGGCGCGGAAGACGCCGGTGG 267  
Qy 61 AspAspSerSerArgIleIleAenGlySerAspCysAspMetHisThrGlnProTrp 80  
Db 268 GATGACAGCAGCAGCCCATCATGATGGATCGGACTGCGATATGCACACCCAGCGTGG 327  
Qy 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
Db 328 CAGCCCGCGTGTGCTPAAGCCCAACCACTCTACTCGCGGGGGTGTGGTGCATCCA 387  
Qy 101 GlnTrpLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120  
Db 388 CAGTGGCTGTACGGCGCCCACTGCAGGAAGAAGTTTTCAGAGTCCGTCTCGGCCAC 447  
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
Db 448 TACTCCCTGTACACAGTTTATGATTCGGCAGCAGATGTTCCAGGGGGTCAATCCATC 507  
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
Db 508 CCCACCTGGCTACTCCACCTGGCCACTCTAACGACCTCATGCTCATCAACATGAC 567  
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
Db 568 AGAAGAATTCGTCCCACTAAAGATGTGACACCCATCAACGTCTCTCTCATTTGTCCTCT 627  
Qy 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
Db 628 GCTGGGACAAGTCTTGGTGTCTGGCTGGGGGACAACCAAGAGCCCAAGTGCACTTC 687  
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
Db 688 CCTAAGTCTCCAGTCTTGAATATCAGCGTGTCTAAGTCAGAAAAGTGGAGGATGCT 747  
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
Db 748 TACCCGAGACAGATAGATGACACCATGTTCTGCGCCGCGTGACAAAGCAGGTAGACTCC 807  
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
Db 808 TGGCAGGGTGATTTGGGGGGCCCTGTGGTCTGCAATGGCTCCCTGCGAGGACTGCTGCTCC 867  
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
Db 868 TGGGAGATATACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAG 927  
Qy 281 PheThrLysTrpIleGlnGlnThrIleGlnAlaAsnSer 293  
Db 928 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 966

RESULT 4  
AAZ06259

ID AAZ06259 standard; DNA; 1381 BP.

XX AAZ06259;

XX 30-SEP-1999 (first entry)

XX Human secreted protein gene No. 27.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

XX WO9935158-A1.

XX 15-JUL-1999.

XX 06-JAN-1999; 99WO-US00108.

XX 07-JAN-1998; 98US-0070704.

XX 07-JAN-1998; 98US-0070657.

XX 07-JAN-1998; 98US-0070658.

XX 07-JAN-1998; 98US-0070692.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Brewer LA, Duan RD, Ebner R, Lafleur DW, NI J;

XX Olsen HS, Rosen CA, Ruben SM, Soppet DR;

XX WPI; 1999-444190/37.

XX P-PSDB; AAY38426.

XX New isolated human genes and the secreted polypeptides they encode

XX Claim 1; Page 173-174; 227pp; English.

XX This sequence represents a nucleic acid molecule which encodes a  
CC secreted human protein. The gene number is given in the descriptor line.  
CC The gene can be used to generate fusion proteins by linking to the gene  
CC to a human immunoglobulin Fc portion (e.g. AAZ06210) for increasing the  
CC stability of the fused protein as compared to the human protein only.  
CC The invention relates to 36 novel genes and their fragments (nucleic  
CC acid sequences: AAZ06219-206263; amino acid sequences AAY38386-Y38498)  
CC which are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Also, pathological  
CC conditions can be diagnosed by determining the amount of the new  
CC polypeptides in a sample or by determining the presence of mutations in  
CC the new polynucleotides. Specific uses are described for each of the 36  
CC polynucleotides, based on which tissues they are most highly expressed in  
CC (see AAZ06219 for described uses).

XX SQ Sequence 1381 BP; 329 A; 425 C; 353 G; 274 T; 0 other;

## Alignment Scores:

Pred. No.: 1-75e-137 Length: 1381  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-09-936-271B-14 (1-293) x AAZ06259 (1-1381)

Qy 1 MetalThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
Db 129 ATGGCTACAGAACACCCCTGGATGGTGGTGTCTGTCTGTATCAGACCTTGCTT 188



QY	21	LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn	40
Db	189	CTGGGGTACAGAGCATGTCGCCAACAAATGATGTTCTGTGACCAACCCCTCTAAC	248
QY	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer	60
Db	249	ACCGTCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGGAAGACGCCCGGTG	308
QY	61	AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTTP	80
Db	309	GATGACAGCAGCAGCCGATCATCATGATCGGATCGGATGATGATGATGATGATG	368
QY	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro	100
Db	369	CAGGCGCGCTGTGCTAAGGCCCAACACAGCTCTACTGCGGGCGGTGTTGTGATCCA	428
QY	101	GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValargLeuGlyHis	120
Db	429	CAGTGGCTGCTCAGCGCCGCCACTCGAGGAAGAAATTTTCAGAGTCCGCTCGGCCAC	488
QY	121	TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle	140
Db	489	TACTCCCTGTACCACTTTATGATCTGGCAGCAGATGTTCCAGGGGTCAATCCATC	548
QY	141	ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn	160
Db	549	CCCCACCCCTGGCTACTCCACCCCTGACACTCTAACGACCTCATGCTCATCAACTGAAC	608
QY	161	ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer	180
Db	609	AGAAGAATTCGTCCTCCACTAAGATGTCAGACCCATCAAGCTCTCTCTCTCTCTCT	668
QY	181	AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe	200
Db	669	GCTGGGCAAAAGTCTGTGTGTCTGCTGGGGCAACCAAGAGCCCCCAAGTGCATTC	728
QY	201	ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla	220
Db	729	CCTAAGTCTCTCCAGTCTTGAATATCAGCGTCTAAGTCAGAAAAGGTGCGAGGATGCT	788
QY	221	TyrProArgGlnIleAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer	240
Db	789	TACCCGAGACAGATAGATACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCC	848
QY	241	CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer	260
Db	849	TGCCAGGGTGATCTCGGGGGCTGTGTCTGCAATGGCTCCCTGCGAGGACTCGTGTC	908
QY	261	TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys	280
Db	909	TGGGAGATTACCCCTGTGCGCGGCCAACAGACCGGGTGTCTACAGAACCTCTGCAAG	968
QY	281	PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer	293
Db	969	TTCCACCAAGTGGATCCAGGAACCACTCCAGGCCAATCC	1007
RESULT 5			
AA60578			
ID	AA60578 standard; DNA; 1499 BP.		
XX			
AC	AA60578;		
XX			
DT	02-AUG-1999 (first entry)		
XX	Human keratinocyte derived protease (KDP) encoding DNA.		
DE			
XX	Keratinocyte derived protease; KDP; protease; keratinocyte; human;		
KW	skin care product; skin flaking; dandruff; laundry detergent composition;		
KW	cleaning composition; dishwashing product; ss.		
OS	Homo sapiens.		
XX			
FH	Key Location/Qualifiers		

FT	5' UTR	1..290	
FT		/tag- a	
FT	CDS	291..1172	
FT		/tag- b	"keratinocyte derived protease"
FT		/product-	
FT	sig_peptide	291..488	
FT		/tag- c	
FT	mat_peptide	489..1169	
FT		/tag- d	
FT	3' UTR	1173..1499	
FT		/tag- e	
XX			
XX	WO9918219-A1.		
XX			
XX	15-APR-1999.		
XX			
XX	03-OCT-1997;	97WO-US17864.	
XX			
XX	03-OCT-1997;	97WO-US17864.	
XX			
XX	(PROC ) PROCTER & GAMBLE CO.		
XX			
XX	Kitado H, Yoshikawa A, Zaiki T;		
XX			
XX	WPI; 1999-287737/24.		
XX	P-PSDB; AAY16777.		
XX			
XX	Novel human protease useful for treating or preventing skin flaking		
XX			
XX	Claim 3; Page 32-34; 43pp; English.		
XX			
XX	This DNA encodes a human keratinocyte derived protease (KDP) polypeptide.		
XX	The protease is substantially similar to a sequence encoded by the insert		
XX	in plasmid FERM BP-6129. The KDP polypeptide, particularly to treat or		
XX	is specifically used in skin care products, particularly to treat or		
XX	prevent skin flaking, e.g. dandruff, and in laundry detergent		
XX	compositions. More generally the KDP polypeptide can be used in any		
XX	cleaning composition, e.g. hard surface or dishwashing products.		
XX			
XX	Sequence 1499 BP; 310 A; 469 C; 398 G; 322 T; 0 other;		
Alignment Scores:			
Pred. No.:	1..95e-137	Length:	1499
Score:	1608.00	Matches:	293
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0
US-09-936-271B-14 (1-293) x AAX60578 (1-1499)			
QY	1	MetAlaThrAlaArgProTTPMetTTPValLeuCysAlaLeuIleThrAlaLeuLeu	20
Db	291	ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGTGCTGTGATCAGCGCTTGGCT	350
QY	21	LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn	40
Db	351	CTGGGGTCTACAGAGCATGTTCTCGCCAAACATGATGTTCTGTGACCAACCCCTTAAC	410
QY	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer	60
Db	411	ACCGTGCCTCTGGGAGCAACAGGACCTGGGAGCTGGGCGCGGGGAAGACCGCGTGC	470
QY	61	AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTTP	80
Db	471	GATGACAGCAGCGCGCATCATCAATGATCGGATGCGATGATGATGATGATGATG	530
QY	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro	100
Db	531	CAGGCGCGCTGTGTGTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGTGTCATCCA	590
QY	101	GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValargLeuGlyHis	120

```
Db 591 CAGTGGCTGCTCAGCGCGCCGCGCCACTGCAGGAAGAAAGTTTTCAGAGTCCCTCTCGGGCCAC 650
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 651 TACTCCCTGTGTCACCAAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGGTCAATCCATC 710
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160
Db 711 CCCACCCCTGGCTACTCCCACTGGCCACTTAACGACCTCATGCTCATCAAACTGAAC 770
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 771 AGAAGAATCGTCCCACTAAAGATGTCAGACCCATCAACGCTCCTCTCATTTGCCCTCT 830
Qy 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
Db 831 GCTGGGACAAAGTCTGGTGTCTGGCTGGGGACACCAAGAGCCGCCCAAGTGCACCTC 890
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 891 CCTAAGCTCTCCAGTCTTGAATATCAGCGTGTCTAAGTCAGAAAGGTGCGAGGATGCT 950
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 951 TACCCGAGACAGATAGATGACACCATGTTCTGCCCGGTGACAAAGCAGGTAGAGATCC 1010
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db 1011 TGCCAGGTGATTTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1070
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 1071 TGGGGAGATTACCCCTTGTGCCCGGCCCAACAGACCGGGGTGTCTACACGAACCTCTGCAAG 1130
Qy 281 PheThrLysTyrPileGlnGluThrIleGlnAlaAsnSer 293
Db 1131 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1169
RESULT 6
AAZ06245
ID AAZ06245 standard; DNA; 1516 BP.
XX AC AAZ06245;
XX DT 30-SEP-1999 (first entry)
XX DE Human secreted protein gene No. 27.
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX OS Homo sapiens.
XX PN W09J35158-A1.
XX PD 15-JUL-1999.
XX PF 06-JAN-1999; 99WO-US00108.
XX PR 07-JAN-1998; 98US-0070704.
XX PR 07-JAN-1998; 98US-0070657.
XX PR 07-JAN-1998; 98US-0070658.
XX PR 07-JAN-1998; 98US-0070692.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;
```

```
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;
XX WPI: 1999-444190/37.
DR P-PSDB; AAY38412.
XX PT New isolated human genes and the secreted polypeptides they encode
XX Claim 1; Page 163-164; 227pp; English.
PS This sequence represents a nucleic acid molecule which encodes a
XX secreted human protein. The gene number is given in the descriptor line.
CC The gene can be used to generate fusion proteins by linking to the gene
CC to a human immunoglobulin Fc portion (e.g. AAZ06210) for increasing the
CC stability of the fused protein as compared to the human proteins only.
CC The invention relates to 36 novel genes and their fragments (nucleic
CC acid sequences: AAZ06219-206263; amino acid sequences AAY38386-Y38498)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 36
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAZ06219 for described uses).
XX SQ Sequence 1516 BP; 338 A; 472 C; 381 G; 324 T; 1 other;
Alignment Scores:
Pred. No.: 1,98e-137 Length: 1516
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0
US-09-936-271B-14 (1-293) x AAZ06245 (1-1516)
Qy 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
Db 254 ATGGCTACAGCAAGACCCCTGGATGGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 313
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 314 CTGGGGTTCACAGAGCATGTTCTCGCCCAACAATGATGTTTCTCTGTGACCACTCTTAC 373
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 374 ACCGTGCCCTCTGGGAGCAACAGGACCTGGAGCTGGGGCGGGGGAAGACGCCGGTCG 433
Qy 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMethHisThrGlnProTrp 80
Db 434 GATGACAGCAGCAGCCCATCATCAATGGATCGGACTATGCGATATGCGACCCAGCCGTGG 493
Qy 81 GlnAlaAlaLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 494 CAGCCCGCGTGTGTCTAAGGCCCAACCACTCTACTGCGGGGGGGTGTGTGTGTGTGTGT 553
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
Db 554 CAGTGGCTGCTCAGCGCGCCCACTGCAGGAAGAAGTTTTCAGAGTCCCTCTCGGGCCAC 613
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 614 TACTCCCTGTGACAGATTTATGAATCTGGGCAGCAGATGTTCCAGGGGGTCAATCCATC 673
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160
Db 674 CCCACCCCTGGCTACTCCCACTGGCCACTCTTAACGACCTCATGCTCATCAAACTGAAC 733
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 734 AGAAGAATTCGTCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCTCATTTGCCCTCT 793
Qy 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
```

|||||  
794 GCTGGACAAAGTGTCTGGTGGCTGGGGCAACCAAGAGCCCCCAAGTGCATCTC 853  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnIysArgCysGluAspAla 220  
Db 854 CTAAGTGCTCTCAGTGGCTTGAATATCAACGGTGTAAAGTCAGAAAGGTGCGAGATGCT 913  
QY 221 TyrProArgGlnIleAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
Db 914 TACCCGAGACAGATAGATCACACCATGTTCTGGCGGGTGACAAAGCAGGTAGACTCC 973  
QY 241 CysGlnGlyAspSerGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
Db 974 TGCCAGGGTGAATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1033  
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
Db 1034 TGGGGAGATTACCTTGTCCCGGGCCCAACAGACGGGTGTCTACAGAACCTCTGCAAG 1093  
QY 281 PheThrLysTrpIleGlnThrIleGlnAlaAsnSer 293  
Db 1094 TTCACCAAGTGGATCCAGGAACCATCCAGGCCAATCC 1132  
RESULT 7  
AAZ23317  
ID AAZ23317 standard; DNA; 1539 BP.  
XX  
AC AAZ23317;  
XX  
DT 03-DEC-1999 (first entry)  
XX  
DE Human PDSP-1 DNA.  
XX  
KW PDSP-1; Tango 114; prostate-derived serine protease; fusion protein;  
KW treatment; proliferative disorder; prostate cancer; screening;  
KW predictive medicine; growth factor; growth factor biosynthesis;  
KW cellular proliferation; growth factor binding protein; metastasis;  
KW cellular differentiation; prostate development; detection; forensic;  
KW serine protease; human; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 244..1125  
FT FT /\*tag= a  
FT FT /product= "PDSP-1"  
XX  
PN W09946391-A2.  
XX  
PD 16-SEP-1999.  
XX  
PF 12-MAR-1999; 99WO-0505416.  
XX  
PR 12-MAR-1998; 98US-0041400.  
XX  
PA (MILL-) MILLENNIUM BIOTHEAPEUTICS INC.  
XX  
PI Holtzman DA;  
XX  
DR WPI; 1999-561678/47.  
DR P-PSDB; AAY30524.  
XX  
PT Novel prostate derived serine protease polynucleotides and polypeptides  
PT used to modulate cellular processes  
XX  
PS Claim 1a; Fig 1A; 94pp; English.  
XX  
CC This invention describes a novel human prostate-derived serine protease  
CC (PDSP) polynucleotide (also known as Tango 114). The prostate derived  
CC serine protease (PDSP) polypeptide is used to modulate a variety of  
CC cellular processes. It can be used to produce fusion proteins. PDSP is used  
CC to treat proliferative disorders, e.g. prostate cancer. The protein may  
CC also be used to produce antibodies, and to identify antagonists and

agonists. The PDSP polynucleotides, polypeptides, homologs and antibodies  
can be used in screening assays; predictive medicine; and methods of  
treatment. PDSP cleaves growth factors, and can be used for the  
modulation of growth factor biosynthesis; generation of active peptides;  
regulation of cellular proliferation; degradation of growth factor  
binding proteins; regulation of cellular differentiation; regulation of  
metastasis; and regulation of prostate development. The PDSP  
polynucleotides can be used to express the protein; to detect PDSP mRNA;  
to detect genetic alterations in the PDSP gene; in forensic biology; and  
as a source of primers and probes. As serine protease have important  
roles in cellular processes, there exists a need for identifying novel  
serine protease such as the prostate derived serine protease (PDSP) of  
the invention. This sequence encodes the human PDSP-1 described in the  
invention.  
XX  
SQ Sequence 1539 BP; 340 A; 470 C; 413 G; 315 T; 1 other;  
Alignment Scores:  
Pred. No.: 2,02e-137 Length: 1539  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0  
US-09-936-271B-14 (1-293) x AAZ23317 (1-1539)  
QY 1 MetAlaThrAlaArgProThrProMetTrpValLeuCysAlaLeuIleThrAlaLeu 20  
Db 244 ATGGCTACAGCAAGACCCCTGGATGTGGGTGCTGTGCTGTATCATCAGCCCTTGCTT 303  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCysAspHisProSerAsn 40  
Db 304 CTGGGGGTACAGAGCATGTTCTCGCCACATGATGTTTCTGTGACACCCCTCTAAC 363  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
Db 364 ACCGTGCCCTCTGGGAGCAACACGAGACCTGGGAGCTGGGCGGGGAGAGCCCGGTG 423  
QY 61 AspAspSerSerArgIleLeuAsnGlySerAspCysAspMethIsthrGlnProTrp 80  
Db 424 GATGACAGCAGCGCGCATCATCATGATCCGCTGCGATGATGACACCCAGCGGTG 483  
QY 81 GlnAlaAlaLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
Db 484 CAGCGCGGCTGTGCTAAGGCCCAACACGACTCTACTGCGGGCGGTGTTGGTCATCCA 543  
QY 101 GlnTrpLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
Db 544 CAGTGGCTGCTCAGCGCGCCCACTGCAGGAAGAAAGTTTTCAGAGTCCGCTCTCGGCCAC 603  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
Db 604 TACTCCCTGTCCACAGTTTATGAATCTGGGAGCAGATGTTCCAGGGGTCAATCATC 663  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
Db 664 CCCCACCCCTGGCTACTCCACCCCTGACACTCTAAGGACCTCATGCTCAACTGAAC 723  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
Db 724 AGAAGAATTCGTCCTCCACTAAAGATGTACAGACCATCAACGTCTCTCTCATTTGCTCCT 783  
QY 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
Db 784 GCTGGGACAAAGTGTGTGTGCTGGGGGACCAACCAAGAGCCCCCAAGTGCATCTTC 843  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
Db 844 CTAAGTGCTCTCAGTGGCTTGAATATCAGCGTGTAAAGTCAGAAAGGTGCGAGATGCT 903  
QY 221 TyrProArgGlnIleAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
|||||

```
Db 904 TACCCGACAGATAGATGACACACATGTTCTGCGCCGGTGACAAAGCAGGTAGAGACTCC 963
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
    |||||
Db 964 TGCAGGGTGATTCCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGACTCGTGCTCC 1023
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
    |||||
Db 1024 TGGGAGATTACCTTCTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAG 1083
Qy 281 PheThrIysTrpIleGlnGluThrIleGlnAlaAsnSer 293
    |||||
Db 1084 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1122

RESULT 8
AAZ65070
ID AAZ65070 standard; cDNA; 1570 BP.
XX
AC AAZ65070;
XX
DT 05-APR-2000 (first entry)
XX
DE Membrane-bound protein PRO1132 encoding cDNA.
XX
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunoadhesin; gene mapping; ss.
OS Homo sapiens.
XX
PN WO9963088-A2.
XX
PD 09-DEC-1999.
XX
PF 02-JUN-1999; 99WO-US12252.
XX
PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 04-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.
PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089952.
PR 19-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 23-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 07-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
```

PR 17-AUG-1998; 98US-0096891.  
 PR 17-AUG-1998; 98US-0096894.  
 PR 17-AUG-1998; 98US-0096895.  
 PR 17-AUG-1998; 98US-0096897.  
 PR 18-AUG-1998; 98US-0096949.  
 PR 18-AUG-1998; 98US-0096950.  
 PR 18-AUG-1998; 98US-0096959.  
 PR 18-AUG-1998; 98US-0096960.  
 PR 18-AUG-1998; 98US-0097022.  
 PR 19-AUG-1998; 98US-0097141.  
 PR 20-AUG-1998; 98US-0097218.  
 PR 24-AUG-1998; 98US-0097661.  
 PR 26-AUG-1998; 98US-0097951.  
 PR 26-AUG-1998; 98US-0097952.  
 PR 26-AUG-1998; 98US-0097954.  
 PR 26-AUG-1998; 98US-0097955.  
 PR 26-AUG-1998; 98US-0097971.  
 PR 26-AUG-1998; 98US-0097974.  
 PR 26-AUG-1998; 98US-0097978.  
 PR 26-AUG-1998; 98US-0097979.  
 PR 26-AUG-1998; 98US-0097986.  
 PR 26-AUG-1998; 98US-0098014.  
 PR 31-AUG-1998; 98US-0098525.  
 PR 16-SEP-1998; 98US-0100634.  
 PR 12-JAN-1999; 99US-0115565.  
 XX  
 PA (GETH ) GENENTECH INC.

PI Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;  
 PI Wood WI, Yuan J;  
 XX  
 XX  
 DR WPI; 2000-072883/06.  
 DR P-PSDB; AAY66726.

PT Membrane-bound proteins and related nucleotide sequences -  
 XX  
 PS Claim 2; Fig 225; 822pp; English.

XX The invention provides membrane-bound PRO polypeptides and  
 CC polynucleotides encoding them. The PRO sequences of the invention were  
 CC identified based on extracellular domain homology screening. The PRO  
 CC sequences have homology with proteins including LDL receptors, TIE  
 CC ligands and various enzymes. The membrane-bound proteins and receptor  
 CC molecules are useful as pharmaceutical and diagnostic agents. Receptor  
 CC immunoadhesins, for instance, can be used as therapeutic agents to block  
 CC receptor-ligand interactions. The membrane-bound proteins can also be  
 CC employed for screening of potential peptide or small molecule inhibitors  
 CC of the relevant receptor/ligand interaction. The PRO encoding sequences  
 CC are useful as hybridization probes, in chromosome and gene mapping and in  
 CC the generation of antisense RNA and DNA. PRO nucleic acid sequences  
 CC will also be useful for the preparation of PRO polypeptides, especially  
 CC by recombinant techniques.

XX SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

Alignment Scores:  
 Pred. No.: 2,07e-137 Length: 1570  
 Score: 1608.00 Matches: 293  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

US-09-936-271B-14 (1-293) x AAZ65070 (1-1570)

QY 1 MetalAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
 DB 354 ATGGCTACAGCAAGACCCCTGGATGGTGGTGGTCTGTGTCTGATCAGCCITGCTT 413  
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
 DB 414 CTGGGGGTACAGAGCATGTTCTGCCCAACAATGATGTTTCTGTGACCAACCCCTTAAC 473

QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
 DB ACCGTGCCCTCTGGGAGCAACACGAGACCTGGGAGCTGGGGCGGGGAAACACGCCCGGTG 533  
 QY 61 AspAspSerSerArgIleLeuAsnGlySerAspCysAspMetHisThrGlnProTrp 80  
 DB GATGACAGCAGCAGCGCATCATCAATGATCGGACTCGGATATGCACACCCAGCGGTG 593  
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
 DB CAGGGCGCGCTGTGCTAAGGCCCAACACGACTCTACTGGGGCGGTGTGGTGCATCCA 653  
 QY 101 GlnTrpLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
 DB CAGTGGCTGCTCAGGGCCGCCACTGCAGAGAGAAGATTTTCAGATCCGCTCGTCCGCCAC 713  
 QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
 DB TACTCCCTGTCAACAGTTTATGAATCTGGGACGACAGATGTTCCAGGGGTCAAAATCCATC 773  
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
 DB CCCACCCCTGGTACTCCACCTGGCCACTTAAGACCTCATGCTCATCAACATGAAC 833  
 QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
 DB AGAAGNATTCGTCCTCCACTAAAGATGTCACACCCATCAAGTCTCTCTCATTTGCTCCTCT 893  
 QY 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
 DB GCTGGGACAAAGTGTGTGTGCTGGCTGGGGACAAACCAAGAGCCCCCAAGTGCACATTC 953  
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
 DB CTTAAGGTCTCTCCAGTGTGAATATACGGTGTAAAGTGAAGGAGGTCGAGGATGCT 1013  
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
 DB TACCCGAGACATAGATGACACCATGTTCTGGCGGGTGACAAACAGGATAGAGACTCC 1073  
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
 DB TGCCAGGGTGATTTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTCC 1133  
 QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
 DB TGGGGAGATACCTTGTGCGGGCCCAACAGACCGGGGTCTTACACGAACCTCTCGAAG 1193  
 QY 281 PheThrLysTrpIleGlnThrIleGlnAlaAsnSer 293  
 DB TTCACCAAGTGGATCCAGGAACCATCCAGGCCAATCC 1232  
 DB  
 RESULT 9  
 AAS21471  
 ID AAS21471 standard; cdNA; 1570 BP.  
 XX  
 AC AAS21471;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cdNA sequence encoding for PRO1132 polypeptide.  
 XX  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200140466-A2.  
 XX  
 PD 07-JUN-2001.

PF 01-DEC-2000; 2000WO-US32678.  
 XX  
 PR 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 09-DEC-1999; 99US-0170262.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04114.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 20-MAR-2000; 2000WO-US07377.  
 PR 21-MAR-2000; 2000WO-US07532.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 10-NOV-2000; 2000WO-US30873.  
 XX

(GETH ) GENENTECH INC.

XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX

DR WPI: 2001-408281/43.  
 DR P-PSDB; AAU12399.

XX Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical .

XX Claim 3; Fig 455; 813pp: English.

XX AAS21244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 XX

SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

Alignment Scores:

Pred. No.: 2,07e-137 Length: 1570  
 Score: 1608.00 Matches: 293  
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-936-271B-14 (1-293) x AAS21471 (1-1570)  
 QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
 DB 354 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTCTGTATCAGACCTTGTCTT 413  
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
 DB 414 CTGGGGTCACAGAGATGTCTCGCCAAATGATGTTTCTCTGTGACCACTCTTAAC 473  
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
 DB 474 ACCGTGCTCTGGAGCAACAGGACCTGGAGCTGGGGCCGGGAAGACGCCGGTGG 533  
 QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80  
 DB 534 GATGACAGCAGCAGCCGCATCATCAATGGATCGGACTGCGATATGCACACCCAGCCGTGG 593  
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrcysGlyAlaValLeuValHisPro 100  
 DB 594 CAGCGCGCTGTGTCTAAGGCCCAACAGCTCTACTGCGGGGGGTGTGGTGCATCCA 653  
 QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
 DB 654 CAGTGGCTGCTCAGCGCCGCCACTGAGGAGCAAGAAAGTTTTCAGAGTCCGCTCGGCCAC 713  
 QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
 DB 714 TACTCCCTGTACACAGTTTATGAATCTGGGCAGCAGATGTTCCAGGGGGTCAATCCATC 773  
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
 DB 774 CCCACCTCGCTACTCCCACTGCGCCACTTAACGACCTCATGCTCATCAACTGAAC 833  
 QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
 DB 834 AGAAGATTCTGCCACTAAAGATGTCCAGACCCATCAACGCTCTCTCATTTGCCCTCT 893  
 QY 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
 DB 894 GCTGGGACAAAGTCTTGGTGTCTGGCTGGGGACAAACCAAGAGCCCCCAAGTGCACCTT 953  
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
 DB 954 CCTAAGTCTCTCCAGTCTTGAATATCAGCGTCTAAGTCAAGAAAGTCCGAGGATGCT 1013  
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
 DB 1014 TACCCGAGACAGATAGATGACACCATGTTCTGGCCGGTGACAAAGCAGGTAGAGACTCC 1073  
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
 DB 1074 TGGCAGGGTATCTTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1133  
 QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
 DB 1134 TGGGGAGATTACCTTGTGCGCGCCCAACAGACCCGGGTCTCTACACGACCTCTGCAAG 1193  
 QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
 DB 1194 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1232

RESULT 10

AAF44216

ID AAF44216 standard; cDNA; 1570 BP.

AC AAF44216;

XX

DT 02-APR-2001 (first entry)



ID XX ABK28605 standard; cDNA; 1570 BP.  
AC XX ABK28605;  
DF XX 09-APR-2002 (first entry)  
XX XX Human DNA59767-1489 encoding PRO1132.  
XX XX Human; ss; gene; PRO; antiinflammatory; ophthalmological; vasotropic;  
KW retinal cell injury; ocular disease; retinitis pigmentosa;  
KW macular degeneration; retinal detachment; retinal tear; retinopathy;  
KW acute retinal necrosis syndrome; macular hole; degenerative myopia;  
KW purtscher's retinopathy; oedema; ischaemic condition;  
KW retinal vision occlusion; collagen vascular disease;  
KW thrombocytopaenic purpura; uveitis; retinal vasculitis; Eales disease;  
KW systemic lupus erythematosus; environmental trauma.  
XX OS Homo sapiens.  
XX XX WO200109327-A2.  
XX XX 08-FEB-2001.  
XX XX 28-JUL-2000; 2000WO-US20710.  
XX XX 28-JUL-1999; 99US-146222P.  
PR 13-SEP-1999; 99WO-US20944.  
PR 15-SEP-1999; 99WO-US21090.  
PR 29-NOV-1999; 99WO-US28214.  
PR 30-NOV-1999; 99WO-US28313.  
PR 01-DEC-1999; 99WO-US28301.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
XX XX (GETH ) GENENTECH INC.  
XX XX Ashkenazi AJ, Baker KP, Goddard A, Godowski PJ, Gurney AL;  
PI Kljavin IJ, Lafleur M, Mark MR, Marsters SA, Pitti RM;  
PI Watanabe CK, Wood WI;  
XX WPI; 2002-130120/17.  
DR P-PSDB; AAU81966.  
XX XX Promoting survival of retinal cells, or delaying or preventing retinal  
PT cell injury or death, by contacting retinal cells with PRO175, 220,  
PT 216, 243, 306, 346, 322, 536, 943, 840, 826, 1068 or PRO1132  
PT polypeptide -  
XX Claim 33; Fig 28; 152pp; English.  
XX XX The invention relates to promoting the survival of retinal cells, or  
CC delaying or preventing retinal cell injury or death, by contacting the  
CC retinal cells with the polypeptide such as PRO175, PRO220, PRO216,  
CC PRO243, PRO306, PRO346, PRO332, PRO336, PRO943, PRO840, PRO826,  
CC PRO1068 or PRO1132 polypeptide. Also included are the nucleic acids  
CC encoding the PRO proteins, a vector comprising the nucleic acid, a host  
CC cell comprising the vector, and anti-PRO antibody. The PRO proteins are  
CC useful for promoting survival of retinal cells (retinal neurons such as  
CC retinal ganglion cells, displaced retinal ganglion cells, amacrine  
CC cells, displaced amacrine cells, horizontal neurons or bipolar neurons,  
CC rod photoreceptors, or supportive cells such as Muller cells or pigment  
CC epithelial cells), or delaying or preventing retinal cell injury or  
CC death caused by ocular disease (which is or is associated with  
CC retinitis pigmentosa, macular degeneration, retinal detachment, retinal

CC tear, retinopathy, retinal degenerative disease, macular hole,  
CC degenerative myopia, acute retinal necrosis syndrome, traumatic  
CC chorioretinopathy or contusion, Purtscher's retinopathy, oedema, an  
CC ischaemic condition, central or branch retinal vision occlusion, retinal  
CC collagen vascular disease, thrombocytopaenic purpura, uveitis, retinal  
CC vasculitis, occlusion associated with Eales disease or systemic lupus  
CC erythematosus), retinal injury or environmental trauma. The retinal  
CC cell injury or death is delayed or prevented by substantially not  
CC causing angiogenesis or mitogenesis. The present sequence is a cDNA  
CC encoding a PRO protein.  
XX XX  
SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 2.07e-137 Length: 1570  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0  
  
US-09-936-271B-14 (1-293) x ABK28605 (1-1570)  
  
Qy 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuLeuThrAlaLeuLeu 20  
Db 354 ATGGCTACAGCAGACCCCTGGATGCTGGTGTCTGTCTGTGATCACAGCCTTGCTT 413  
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
Db 414 CTGGGGGTGACAGAGCATGTCTCGCCAAACAATATGTTCTCTGTGACCCCTCTTAA 473  
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
Db 474 ACCGTGCCCTCTGGGAGCAACACGAGCTGGGAGCTGGGGCGGGAGAGACCCCGGTG 533  
Qy 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80  
Db 534 GATGACAGCAGCAGCCGCATCATCAATGGATCCGACTCCGATATGCACACCCCGGTG 593  
Qy 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCysGlyAlaValLeuValHisPro 100  
Db 594 CAGCCCGCGCTGTGCTAAGCCCAACAGCTCTACTCGGGGGGGGTGTGTCATCCCA 653  
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
Db 654 CAGTGGCTGCTACGGCGGCCACTGCAGGAAGAAAGTTTCAGAGTCCGCTCGGCCAC 713  
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
Db 714 TACTCCCTGTACCAGTTTATGAATCTGGGCAGCAGATGTTTCCAGGGGGGTCAAAATCC 773  
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160  
Db 774 CCCACCTCGCTACTCCACCTCCCGCCACTTAAACGACCTCATGCTCATCAAACTGAC 833  
Qy 161 ArgArgIleArgProThrLysAspValArgProfileAsnValSerSerHisCysProSer 180  
Db 834 AGAAGAATTCGTCCCACTAAAGATGTCAGACCCCAACGCTCTCTCTCATTTGTCCTCT 893  
Qy 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrLysSerProGlnValHisPhe 200  
Db 894 GCTGGGACAAAGTCTTGGTGTCTGGTGGGGGACAAACCAAGACCCCAAGTGCATTC 953  
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
Db 954 CCTAAGGTCTCCAGTCTTGAATATCAGCGTGTAGTACAGAAAGGTGCGAGATGCT 1013  
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
Db 1014 TACCCGAGACAGATAGATGACACCATGTTCTGCGCCGCTGACAAAGAGGTAGAGACTCC 1073  
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyValSer 260  
Db 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyValSer 260



Db 1074 TGCCAGGTGATCTGGGGGGCCCTGTGTCTGCAATGGCTCCCTCAGGGACTCGTGCC 1133  
QY 261 TtpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
|||||  
Db 1134 TGGGAGATTACCTTGTGGCGGGCCCAACAGACGGGTGTCTACACGAACCTTCGCAAG 1193  
QY 281 PheThrLysTyrIleGlnGluThrIleGlnAlaAsnSer 293  
|||||  
Db 1194 TTCACCAAGTGTGATCCAGGAACCATCCAGGCGCAACTCC 1232  
RESULT 12  
ACA03830  
ID ACA03830 standard; cDNA; 1570 BP.  
XX AC  
XX ACA03830;  
XX  
XX  
XX 23-MAY-2003 (first entry)  
XX cDNA encoding human PRO polypeptide #228.  
XX Human; PRO polypeptide; secreted and transmembrane protein;  
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;  
KW differentiation; chondrocyte; tumour; genetic disorder;  
KW cytosstatic; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX US2003036180-A1.  
XX  
XX 20-FEB-2003.  
XX  
XX 09-MAY-2002; 2002US-0143114.  
XX  
XX 31-MAR-1997; 97WO-US05230.  
XX 12-JUN-1998; 98WO-US12456.  
XX 14-JUL-1998; 98WO-US14552.  
XX 28-AUG-1998; 98WO-US17888.  
XX 10-SEP-1998; 98WO-US18824.  
XX 14-SEP-1998; 98WO-US19093.  
XX 14-SEP-1998; 98WO-US19094.  
XX 14-SEP-1998; 98WO-US19177.  
XX 16-SEP-1998; 98WO-US19330.  
XX 17-SEP-1998; 98WO-US19437.  
XX 29-OCT-1998; 98WO-US21141.  
XX 29-OCT-1998; 98WO-US22991.  
XX 29-OCT-1998; 98WO-US22992.  
XX 20-NOV-1998; 98WO-US24855.  
XX 01-DEC-1998; 98WO-US25108.  
XX 05-JAN-1999; 99WO-US00106.  
XX 08-MAR-1999; 99WO-US05028.  
XX 10-MAR-1999; 99WO-US05190.  
XX 20-APR-1999; 99WO-US08615.  
XX 14-MAY-1999; 99WO-US10733.  
XX 02-JUN-1999; 99WO-US12252.  
XX 01-SEP-1999; 99WO-US20111.  
XX 08-SEP-1999; 99WO-US20594.  
XX 13-SEP-1999; 99WO-US20944.  
XX 15-SEP-1999; 99WO-US21090.  
XX 15-SEP-1999; 99WO-US21547.  
XX 05-OCT-1999; 99WO-US23089.  
XX 29-NOV-1999; 99WO-US28214.  
XX 30-NOV-1999; 99WO-US28313.  
XX 30-NOV-1999; 99WO-US28409.  
XX 01-DEC-1999; 99WO-US28301.  
XX 01-DEC-1999; 99WO-US28634.  
XX 02-DEC-1999; 99WO-US28551.  
XX 02-DEC-1999; 99WO-US28564.  
XX 02-DEC-1999; 99WO-US28565.  
XX 16-DEC-1999; 99WO-US30095.  
XX 20-DEC-1999; 99WO-US30911.  
XX 20-DEC-1999; 99WO-US30999.  
XX 22-DEC-1999; 99WO-US30720.  
XX 30-DEC-1999; 99WO-US31243.

PR 30-DEC-1999; 99WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05746.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 10-MAR-2000; 2000WO-US06319.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 11-AUG-2000; 2000WO-US22031.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 01-MAR-2001; 2001WO-US06666.  
PR 25-MAY-2001; 2001WO-US17092.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 22-JUN-2001; 2001WO-US20116.  
PR 29-JUN-2001; 2001WO-US21066.  
PR 09-JUL-2001; 2001WO-US21735.  
PR 20-DEC-2000; 2000US-0747259.  
PR 28-FEB-2001; 2001US-0796498.  
PR 09-MAR-2001; 2001US-0802706.  
PR 14-MAR-2001; 2001US-0808689.  
PR 22-MAR-2001; 2001US-0816744.  
PR 05-APR-2001; 2001US-0828366.  
PR 10-MAY-2001; 2001US-0854208.  
PR 18-MAY-2001; 2001US-0860216.  
PR 25-MAY-2001; 2001US-0866028.  
PR 25-MAY-2001; 2001US-0866034.  
PR 01-JUN-2001; 2001US-0872035.  
PR 05-JUN-2001; 2001US-0874503.  
PR 14-JUN-2001; 2001US-0882636.  
PR 19-JUN-2001; 2001US-0886342.  
PR 21-JUN-2001; 2001US-0887879.  
PR 18-JUL-2001; 2001US-0908827.  
PR 06-AUG-2001; 2001US-0924419.  
PR 09-AUG-2001; 2001US-0927796.  
PR 16-AUG-2001; 2001US-0931836.  
PR 19-DEC-2001; 2001US-0028072.  
XX  
XX  
XX (GETH ) GENENTECH INC.  
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX WPI: 2003-332040/31.  
XX P-PSDB; ABU66797.  
XX  
XX New secreted and transmembrane PRO nucleic acids, useful for gene  
PT therapy, in chromosome and gene mapping, as chromosome markers, in  
PT tissue typing, and in chromosome identification  
XX

PS Claim 2; Fig 455; 660pp; English.

xx The present invention relates to the isolation of novel human PRO polypeptides, and the polynucleotide sequences encoding them. The PRO polypeptides are secreted and transmembrane proteins. The PRO polypeptides are useful for detecting other PRO polypeptides, for linking bioactive molecules to cells expressing PRO polypeptides, for modulating biological activities of cells expressing PRO polypeptides, and for identifying agonists or antagonists.

CC The PRO polypeptides are useful for stimulating the release of tumour necrosis factor (TNF)-alpha from human blood, for stimulating the proliferation or differentiation of chondrocytes, and detecting the presence of tumours. The polynucleotide sequences encoding PRO polypeptides are useful as hybridisation probes, in chromosome and gene mapping, in the generation of antisense RNA and DNA, in the preparation of PRO polypeptides, for generating transgenic animals or knockout animals, for the genetic analysis of individuals with genetic disorders, and in gene therapy. ACA03603-ACA03877 represent cDNAs encoding the human PRO polypeptides of the invention.

CC Note: The sequence data for this patent was obtained in electronic format directly from the USPTO web site at [seqdata.uspto.gov/psipsdIDentry.html](http://seqdata.uspto.gov/psipsdIDentry.html).

xx Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

#### Alignment Scores:

Pred. No.:	2.07e-137	Length:	1570
Score:	1608.00	Matches:	293
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	25	Gaps:	0

US-09-936-271B-14 (1-293) x ACA03830 (1-1570)

QY	1	MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuThrAlaLeuLeu 20
DB	354	ATGGCTACACAGACCCCTGGATGGTGGTCTCTGTCTGATCAGCGCTTGCCT 413
QY	21	LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
DB	414	CTGGGGGTACACAGCATGTTCTCGCAACAATGATGTTTCTGTGACCACTCTAAC 473
QY	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60
DB	474	ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGCGGGGAGACGCCGCTCG 533
QY	61	AspAspSerSerArgIleIleAsnGlySerAspCysAspMethHisThrGlnProTrp 80
DB	534	GATGACAGCAGCAGCCGATCATATGGATCCGACTGCGATATGCACCCAGCCGTGG 593
QY	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
DB	594	CAGGCGCGCTGTTGTAAGGCCCAACCACTCTACTCGGGCGGTGTGTGTCATCCA 653
QY	101	GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120
DB	654	CAGTGGCTGTCTACGGCGGCCCACTGCAGAGAAAGTTTTCAGAGTCCGCTCTCGGCCAC 713
QY	121	TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
DB	714	TACTCCTGTACACAGTTATGATCTGGGCACAGATGTTCAGGGGGTCAATCCATC 773
QY	141	ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAsn 160
DB	774	CCCCACCTGGCTACTCCACCTCGGCCACTCTAACGACCTCATGTCTCAACATGAAC 833
QY	161	ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
DB	834	AGAAGAATTTCGTCCTCCCACTAAAGATGTTCAGACCCCACTCAACGCTCTCTCTCTCTCT 893
QY	181	AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200

DB	894	GCTGGGACAAAGTGTCTTGGTGTCTGGCTGGGGGACAAACCAAGAGCCCCCAAGTGCACCTTC 953
QY	201	ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
DB	954	CCTAAGGTCTCTCAGTGTGTGATATACAGCGTCTAAGTCAGAAAGGTCCGAGGATGCT 1013
QY	221	TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
DB	1014	TACCCGAGACAGATAGATGACACCATGTTCTGCGCGGTGCACAAAGCAGGTAGAGACTCC 1073
QY	241	CysGlnClyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB	1074	TGCCAGGGTGATTTCTGGGGGGCTGTGGTGTGCAATGGCTCCCTGCAGGACTCGTGTCC 1133
QY	261	TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
DB	1134	TGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGTCTCTACAGCAACCTCTGCAAG 1193
QY	281	PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
DB	1194	TTCCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1232

RESULT 13  
ACA04251 standard; cDNA; 1570 BP.

XX	ACA04251;
XX	27-MAY-2003 (first entry)
XX	Human cDNA encoding a secreted/transmembrane protein, SEQ ID 455.
XX	Human; ss; gene; secreted protein; transmembrane protein; PRO; inflammatory disease; organ failure; atherosclerosis; cardiac injury; infertility; birth defects; premature aging; AIDS; biosensor; acquired immunodeficiency syndrome; cancer; diabetic complication; bioreactor; tumour.
OS	Homo sapiens.
XX	US2003032155-A1.
XX	13-FEB-2003.
XX	03-MAY-2002; 2002US-0137865.
XX	31-MAR-1997; 97WO-US05230.
PR	12-JUN-1998; 98WO-US12456.
PR	14-JUL-1998; 98WO-US14552.
PR	28-AUG-1998; 98WO-US17888.
PR	10-SEP-1998; 98WO-US18824.
PR	14-SEP-1998; 98WO-US19093.
PR	14-SEP-1998; 98WO-US19094.
PR	14-SEP-1998; 98WO-US19177.
PR	16-SEP-1998; 98WO-US19330.
PR	17-SEP-1998; 98WO-US19437.
PR	07-OCT-1998; 98WO-US21141.
PR	29-OCT-1998; 98WO-US22991.
PR	29-OCT-1998; 98WO-US22992.
PR	20-NOV-1998; 98WO-US24855.
PR	01-DEC-1998; 98WO-US25108.
PR	03-JAN-1999; 99WO-US00106.
PR	08-MAR-1999; 99WO-US05028.
PR	10-MAR-1999; 99WO-US05190.
PR	20-APR-1999; 99WO-US08615.
PR	14-MAY-1999; 99WO-US10733.
PR	02-JUN-1999; 99WO-US12252.
PR	01-SEP-1999; 99WO-US20111.
PR	08-SEP-1999; 99WO-US20594.
PR	13-SEP-1999; 99WO-US20944.
PR	15-SEP-1999; 99WO-US21090.
PR	15-SEP-1999; 99WO-US21547.
PR	03-OCT-1999; 99WO-US23089.

PR 29-NOV-1999; 99WO-US28214.  
PR 30-NOV-1999; 99WO-US28313.  
PR 01-DEC-1999; 99WO-US28409.  
PR 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28564.  
PR 02-DEC-1999; 99WO-US28565.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 20-DEC-1999; 99WO-US30999.  
PR 22-DEC-1999; 99WO-US30720.  
PR 30-DEC-1999; 99WO-US31243.  
PR 30-DEC-1999; 99WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05746.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 10-MAR-2000; 2000WO-US06319.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 11-AUG-2000; 2000WO-US22031.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 10-NOV-2000; 2000WO-US30973.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 01-MAY-2001; 2001WO-US06666.  
PR 25-MAY-2001; 2001WO-US17092.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 22-JUN-2001; 2001WO-US20116.  
PR 29-JUN-2001; 2001WO-US21066.  
PR 09-JUL-2001; 2001WO-US21735.  
PR 20-DEC-2000; 2000US-0747259.  
PR 28-FEB-2001; 2001US-0796498.  
PR 09-MAR-2001; 2001US-0802706.  
PR 14-MAR-2001; 2001US-0808689.  
PR 22-MAR-2001; 2001US-0816744.  
PR 05-APR-2001; 2001US-0828366.  
PR 10-MAY-2001; 2001US-0854208.  
PR 18-MAY-2001; 2001US-0854280.  
PR 18-MAY-2001; 2001US-0860216.  
PR 25-MAY-2001; 2001US-0866028.  
PR 25-MAY-2001; 2001US-0866034.  
PR 01-JUN-2001; 2001US-0872035.  
PR 05-JUN-2001; 2001US-0874503.  
PR 14-JUN-2001; 2001US-0882636.  
PR 19-JUN-2001; 2001US-0886342.  
PR 21-JUN-2001; 2001US-0887879.  
PR 18-JUL-2001; 2001US-0908827.  
PR 06-AUG-2001; 2001US-0924419.  
PR 09-AUG-2001; 2001US-0927796.  
PR 16-AUG-2001; 2001US-0931836.  
PR 19-DEC-2001; 2001US-0028072.  
XX

PA (GETH ) GENENTECH INC.  
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski RJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX WPI; 2003-331925/31.  
DR P-PSDB; ABU67073.  
XX  
XX New secreted and transmembrane nucleic acids and polypeptides,  
PT designated as PRO, useful for treating inflammation, organ failure,  
PT atherosclerosis, cardiac injury, infertility, birth defects, premature  
PT aging, AIDS, or cancer  
XX  
PS Claim 2; Fig 455; 659pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising, or which is  
CC at least 80% identical to, or the full-length coding sequence of, any of  
CC the 275 nucleotide sequences, encoding the corresponding PRO polypeptide  
CC (one of 275 secreted or transmembrane proteins). The nucleic acid  
CC further comprises the full-length coding sequence of the DNA deposited  
CC under American Type Culture Collection (ATCC) accession number in a list  
CC given in the specification. Also included are vectors and host  
CC cells for producing PRO proteins, PRO fusion proteins, anti-PRO  
CC antibodies, PRO extracellular domains and mature sequences, methods  
CC of detecting PRO proteins, methods for stimulating the release of  
CC TNF-alpha (tumour necrosis factor alpha) from human blood.  
CC (and the proliferation of differentiation of chondrocyte cells, the  
CC proliferation of, or gene expression in pericyte cells, the release or  
CC proteoglycans from cartilage, proliferation of inner ear utricular  
CC supporting cells, the proliferation of T-lymphocyte cells, the release  
CC of a cytokine from peripheral blood mononuclear cells (PBMC), or the  
CC proliferation of endothelial cells), a method for modulating the uptake  
CC of glucose or free fatty acid (FFA) by skeletal muscle cells.  
CC a method for inhibiting the binding of A-peptide to factor VIIa,  
CC or the differentiation of adipocyte cells, a method for detecting the  
CC presence of a tumour in a mammal and an oligonucleotide probe derived  
CC from any of the nucleotide sequences cited above. The nucleic acids and  
CC polypeptides are useful for treating inflammatory diseases, organ  
CC failure, atherosclerosis, cardiac injury, infertility, birth defects,  
CC premature aging, AIDS (acquired immunodeficiency syndrome), cancer, or  
CC diabetic complications. The nucleic acids are useful as hybridisation  
CC probes, in chromosome and gene mapping, and in generating antisense RNA  
CC or DNA. The polypeptides are useful as pharmaceuticals, diagnostics,  
CC biosensors or bioreactors. Both are useful in tissue typing.  
CC The present sequence encodes a PRO protein of the invention.  
XX  
SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 2,07e-137 Length: 1570  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0  
  
US-09-936-271B-14 (1-293) x ACA04251 (1-1570)  
  
QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeu 20  
|||||  
Db 354 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 413  
  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
|||||  
Db 414 CTGGGGGTACAGACGATGTTCTGCCAACAAATGATGTTCTGTGTGACACCCCTTAAC 473  
  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
|||||  
Db 474 ACCGTGCCCTCTGGGAGCAACACGACCTGGAGCTGGGGCGGGGAGAACGCCCGGTG 533  
  
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMethHisThrGlnProTrp 80  
|||||

```
Db 534 GATGACAGCAGCGCCGATCATCAATGATCGCATGCGATATGCACACCCAGCGTGG 593
Qy 81 GlnAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 594 CAGCGCGCGTGTGTAGGCCCAACCACTCTACTCGGGCGGTGTGGTGCATCCA 653
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120
Db 654 CAGTGGGTGTCACGGCGCCCACTGCAGGAAGAAAGTTTTCAGAGTCCGTCFCGCCAC 713
Qy 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 714 TACTCCCTGTCCAGGTATTAAGATCTGGCGACAGATGTCCAGGGGGTCAATCCATC 773
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAsn 160
Db 774 CCCACCCCTGGCTACTCCCACTCCGCACTCTAACGACCTCATGCTCATCAACTGAAC 833
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 834 AGAAGAATTGCTCCCACTAAAGATGTGACACCATCAACGCTCTCTCATTTGCTCTCT 893
Qy 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
Db 894 GCTGGGCAAGTCTTGGTGTCTGGCTGGGGGACAAACCAAGAGCCCCCAAGTGCATTC 953
Qy 201 ProlLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 954 CCTAAGTCTCCAGTCTTGAATATACAGGTGCTAAGTCAAGAAAGGTCCGAGGATGCT 1013
Qy 221 TyrProArgGlnIleAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 1014 TACCCGAGACAGATAGATGACACCATGTTCTGGCGCGGTGACAAAGCAGGTAGACTCC 1073
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db 1074 TGGCAGGGTATCTGGGGGGCTGTGGTCTGCAATGGCTCTGACAGGACCTCGGTGCC 1133
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 1134 TGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGGTGCTACAGGAACCTCTGCAAG 1193
Qy 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 1194 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1232

RESULT 14
ABX89368
ID ABX89368 standard; cDNA; 1570 BP.
XX
AC ABX89368;
XX
XX
DT 13-MAY-2003 (first entry)
XX
DE DNA encoding novel secreted and transmembrane protein' PRO1132.
XX
KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
KW cardiac insufficiency disorder; cancer; tumour; immune response;
KW adrenal cortical capillary endothelial growth; c-fos induction;
KW vascular endothelial growth factor inhibition; VEGF inhibition;
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
KW retinal neurons cell survival; rod photoreceptor cell survival;
KW retinal disorder; retinitis pigmentosa; kidney disorder;
KW mammalian kidney mesangial cell proliferation; Berger disease;
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;
KW chondrocyte redifferentiation; sports injury; arthritis; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003017563-A1.
XX
PD 23-JAN-2003.
XX
```

---

```
PF 07-MAY-2002; 2002US-0140808.
XX
PR 31-MAR-1997; 97WO-US05230.
PR 12-JUN-1998; 98WO-US12456.
PR 14-JUL-1998; 98WO-US14552.
PR 28-AUG-1998; 98WO-US17888.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19093.
PR 14-SEP-1998; 98WO-US19094.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 29-OCT-1998; 98WO-US22991.
PR 29-OCT-1998; 98WO-US22992.
PR 20-NOV-1998; 98WO-US24855.
PR 01-DEC-1998; 98WO-US25108.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 20-APR-1999; 99WO-US08615.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28851.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 22-DEC-1999; 99WO-US30720.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05746.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 11-AUG-2000; 2000WO-US22031.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
```

PR 01-MAR-2001; 2001WO-US06666.  
 PR 25-MAY-2001; 2001WO-US17092.  
 PR 01-JUN-2001; 2001WO-US17800.  
 PR 22-JUN-2001; 2001WO-US19692.  
 PR 22-JUN-2001; 2001WO-US20116.  
 PR 29-JUN-2001; 2001WO-US21066.  
 PR 09-JUL-2001; 2001WO-US21735.  
 PR 20-DEC-2000; 2000US-0747259.  
 PR 28-FEB-2001; 2001US-0796498.  
 PR 09-MAR-2001; 2001US-0802706.  
 PR 14-MAR-2001; 2001US-0808689.  
 PR 22-MAR-2001; 2001US-0816744.  
 PR 05-APR-2001; 2001US-0828366.  
 PR 10-MAY-2001; 2001US-0854208.  
 PR 18-MAY-2001; 2001US-0860216.  
 PR 25-MAY-2001; 2001US-0866028.  
 PR 25-MAY-2001; 2001US-0866034.  
 PR 01-JUN-2001; 2001US-0872035.  
 PR 05-JUN-2001; 2001US-0874503.  
 PR 14-JUN-2001; 2001US-0882636.  
 PR 19-JUN-2001; 2001US-0886342.  
 PR 21-JUN-2001; 2001US-0887879.  
 PR 18-JUL-2001; 2001US-0908827.  
 PR 06-AUG-2001; 2001US-0924419.  
 PR 09-AUG-2001; 2001US-0927796.  
 PR 16-AUG-2001; 2001US-0931836.  
 PR 19-DEC-2001; 2001US-0028072.

(GETH ) GENENTECH INC.

XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX WPI; 2003-148238/14.  
 DR P-PSDB; ABU59878.

XX Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346  
 PT and PRO1375, which stimulate proliferation of stimulated T-lymphocytes  
 PT are therapeutically useful for enhancing immune response and in cancer  
 PT treatments

XX Claim 2; Fig 455; 659pp; English.

XX The invention describes an isolated human PRO polypeptide. The PRO  
 CC polypeptides are useful in detecting PRO polypeptides in a sample, in  
 CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and  
 CC in modulating at least one biological activity of a cell expressing a PRO  
 CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus  
 CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186  
 CC stimulate adrenal cortical capillary endothelial growth, and PRO536,  
 CC PRO943, PRO828, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126,  
 CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus  
 CC useful for treating conditions or disorders where angiogenesis would be  
 CC beneficial, e.g. wound healing and antagonist of this polypeptide are  
 CC useful for treating cancerous tumours. PRO812 inhibits vascular  
 CC endothelial growth factor (VEGF) stimulated proliferation of endothelial  
 CC cells and is thus useful for inhibiting endothelial cell growth in  
 CC mammals which would be beneficial in inhibiting tumour growth. PRO826,  
 CC PRO1068, PRO1184, and PRO1375 stimulate proliferation of  
 CC stimulated T-lymphocytes and are therapeutically useful for enhancing  
 CC immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of  
 CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of  
 CC rod photoreceptor cells) and therefore are useful for treating retinal  
 CC disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813  
 CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,  
 CC and therefore are useful for treating kidney disorders associated with  
 CC decreased mesangial cell function such as Berger disease or other  
 CC nephropathies associated with dermatitis, herpeticiformis or Crohn's  
 CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the  
 CC proliferation and/or redifferentiation of chondrocytes in culture and  
 CC are thus useful for treating sports injuries, and arthritis. This

CC sequence encodes a novel human PRO protein.  
 XX  
 SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

# Alignment Scores:

Pred. No.: 2,07e-137 Length: 1570  
 Score: 1608.00 Matches: 293  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 25

US-09-936-271B-14 (1-293) x ABX89368 (1-1570)

QY	1	MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu	20
DB	354	ATGGCTACAGCAAGACCCCTGGATGTGGTGCTCTGTGCTGTGATCAGCGCTGCTT	413
QY	21	LeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCysAspHisProSerAsn	40
DB	414	CTGGGGTTCACAGAGCATGTTCTGCCCAACAATGATGTTCTCTGTGACCACTCTTAAC	473
QY	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer	60
DB	474	ACCGTCCCTCTGGGAGCAACCAAGGACCTGGGAGCTGGGGCCGGGAAGACGCCGCTCG	533
QY	61	AspAspSerSerArgIleIleAsnGlySerAspCysAspMethHisThrGlnProTrp	80
DB	534	GATGACAGCAGCAGCCGATCATCAATGATGATGATGATGATGATGATGATGATGATGAT	593
QY	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro	100
DB	594	CAGGCGCGCTGTGCTAAGGCCCAACCAAGCTCTACTGCGGGCGGTGTGGTGATCCA	653
QY	101	GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis	120
DB	654	CAGTGGCTGCTCAGCGCCGCCACTGCAGGAGAGAAAGTTTTCAGAGTCCGCTCGGCCAC	713
QY	121	TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle	140
DB	714	TACTCCTGTACACAGTTTATGAATCTGGGACAGATGTTCCAGGGGGTCAATCATC	773
QY	141	ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn	160
DB	774	CCCCACCTGGCTACTCCACCTCGCCACTCTAACGACTCTCATCACTCACTCACTCACT	833
QY	161	ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer	180
DB	834	AGAAGAATTCGTCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCTCATTTGCTCT	893
QY	181	AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe	200
DB	894	GCTGGGACAAAGT	953
QY	201	ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla	220
DB	954	CCTAAGTCTCTCCAGT	1013
QY	221	TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer	240
DB	1014	TACCCGAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGACTCC	1073
QY	241	CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer	260
DB	1074	TGCCAGGGGTATTTCTGGGGGCTGTGGTCTTGAATGGCTCTCCCTGCGAGGACTCGTGT	1133
QY	261	TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys	280
DB	1134	TGGGGAGATTACCTTTGTGCCCCGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAG	1193
QY	281	PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer	293
DB	1194	TTCAACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC	1232

RESULT 15  
ABX80318  
ID ABX80318 standard; DNA; 1570 BP.  
XX  
AC ABX80318;  
XX  
DT 28-APR-2003 (first entry)  
XX  
DE Human secreted or transmembrane protein related PCR primer #56.  
XX  
KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;  
KW cardiac insufficiency disorder; cancer; tumour; immune response;  
KW adrenal cortical capillary endothelial growth; c-fos induction;  
KW vascular endothelial growth factor inhibition; VEGF inhibition;  
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;  
KW retinal neurons cell survival; rod photoreceptor cell survival;  
KW retinal disorder; retinitis pigmentosa; kidney disease;  
KW mammalian kidney mesangial cell proliferation; Berger disease;  
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;  
KW chondrocyte redifferentiation; sports injury; arthritis; PCR;  
KW primer; ss.  
XX  
OS Homo sapiens.  
XX  
PN US2002132252-A1.  
XX  
PD 19-SEP-2002.  
XX  
PF 14-NOV-2001; 2001US-0990442.  
XX  
05-NOV-1997; 97WO-US20069.  
PR 16-SEP-1998; 98WO-US19330.  
PR 17-SEP-1998; 98WO-US19437.  
PR 07-OCT-1998; 98WO-US21141.  
PR 01-DEC-1998; 98WO-US25108.  
PR 05-JAN-1999; 99WO-US00106.  
PR 08-MAR-1999; 99WO-US05028.  
PR 02-JUN-1999; 99WO-US12252.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 30-NOV-1999; 99WO-US28313.  
PR 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 06-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 10-MAR-2000; 2000WO-US06319.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 30-MAR-2000; 2000WO-US07377.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 15-MAY-2000; 2000WO-US13358.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 11-AUG-2000; 2000WO-US22031.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 29-JUN-2001; 2001WO-US21066.

PR 09-JUL-2001; 2001WO-US21735.  
PR 16-JUN-1997; 97US-049787P.  
PR 17-OCT-1997; 97US-062250P.  
PR 12-NOV-1997; 97US-065186P.  
PR 13-NOV-1997; 97US-065311P.  
PR 24-NOV-1997; 97US-066770P.  
PR 25-FEB-1998; 98US-075945P.  
PR 20-MAR-1998; 98US-078910P.  
PR 28-APR-1998; 98US-083322P.  
PR 07-MAY-1998; 98US-084600P.  
PR 28-MAY-1998; 98US-087106P.  
PR 02-JUN-1998; 98US-087607P.  
PR 02-JUN-1998; 98US-087609P.  
PR 02-JUN-1998; 98US-087759P.  
PR 03-JUN-1998; 98US-087827P.  
PR 04-JUN-1998; 98US-088021P.  
PR 04-JUN-1998; 98US-088023P.  
PR 04-JUN-1998; 98US-088026P.  
PR 04-JUN-1998; 98US-088028P.  
PR 04-JUN-1998; 98US-088029P.  
PR 04-JUN-1998; 98US-088030P.  
PR 04-JUN-1998; 98US-088033P.  
PR 04-JUN-1998; 98US-088326P.  
PR 05-JUN-1998; 98US-088167P.  
PR 05-JUN-1998; 98US-088202P.  
PR 05-JUN-1998; 98US-088212P.  
PR 05-JUN-1998; 98US-088217P.  
PR 09-JUN-1998; 98US-088655P.  
PR 10-JUN-1998; 98US-088734P.  
PR 10-JUN-1998; 98US-088738P.  
PR 10-JUN-1998; 98US-088742P.  
PR 10-JUN-1998; 98US-088810P.  
PR 10-JUN-1998; 98US-088824P.  
PR 10-JUN-1998; 98US-088826P.  
PR 11-JUN-1998; 98US-088858P.  
PR 11-JUN-1998; 98US-088861P.  
PR 11-JUN-1998; 98US-088876P.  
PR 12-JUN-1998; 98US-089105P.  
PR 16-JUN-1998; 98US-089440P.  
PR 16-JUN-1998; 98US-089512P.  
PR 16-JUN-1998; 98US-089514P.  
PR 17-JUN-1998; 98US-089532P.  
PR 17-JUN-1998; 98US-089538P.  
PR 17-JUN-1998; 98US-089598P.  
PR 17-JUN-1998; 98US-089599P.  
PR 17-JUN-1998; 98US-089600P.  
PR 17-JUN-1998; 98US-089653P.  
PR 18-JUN-1998; 98US-089801P.  
PR 18-JUN-1998; 98US-089907P.  
PR 18-JUN-1998; 98US-089908P.  
PR 28-AUG-2001; 2001US-0941992.

(GETH ) GENENTECH INC.

Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;  
Grimaldi JC, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;  
Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;  
Zhang Z;

WPI; 2003-247083/24.

Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346  
and PRO1375, which stimulate proliferation of stimulated T-lymphocytes  
are therapeutically useful for enhancing immune response and in cancer  
treatments

Example 99; Page 259; 648pp; English.

The invention describes an isolated human PRO polypeptide. The PRO  
polypeptides are useful in detecting PRO polypeptides in a sample, in  
linking a bioactive molecule to a cell expressing a PRO polypeptide, and  
in modulating at least one biological activity of a cell expressing a PRO

CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus  
 CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186  
 CC stimulate adrenal cortical capillary endothelial growth, and PRO536,  
 CC PRO943, PRO828, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126,  
 CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus  
 CC useful for treating conditions or disorders where angiogenesis would be  
 CC beneficial, e.g. wound healing and antagonist of this polypeptide are  
 CC useful for treating cancerous tumours. PRO812 inhibits vascular  
 CC endothelial growth factor (VEGF) stimulated proliferation of endothelial  
 CC cells and is thus useful for inhibiting endothelial cell growth in  
 CC mammals which would be beneficial in inhibiting tumour growth. PRO826,  
 CC PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of  
 CC stimulated T-lymphocytes and are therapeutically useful for enhancing  
 CC immune response. PRO828, PRO826, PRO1068 or PRO132 enhance survival of  
 CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of  
 CC rod photoreceptor cells) and therefore are useful for treating retinal  
 CC disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813  
 CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,  
 CC and therefore are useful for treating kidney disorders associated with  
 CC decreased mesangial cell function such as Berger disease or Crohn's  
 CC nephropathies associated with dermatitis, herpeticiformis or Crohn's  
 CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the  
 CC proliferation and/or redifferentiation of chondrocytes in culture and  
 CC are thus useful for treating sports injuries, and arthritis. This  
 CC sequence represents a primer used in the isolation of DNA encoding  
 CC novel human PRO polypeptides.

XX  
 SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

## Alignment Scores:

Pred. No.:	2,07e-137	Length:	1570
Score:	1608.00	Matches:	293
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	25	Gaps:	0

US-09-936-271B-14 (1-293) x ABX80318 (1-1570)

QY	1	MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu	20
DB	354	ATGGCTACAGCAGACACCCCTGGGATGGGTCTGTGCTGATCACACCTTGTCT	413
QY	21	LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn	40
DB	414	CTGGGGTCCACAGAGCATGTTCTCGCCACACATGATGTTCTGTGACACCCCTTAAC	473
QY	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer	60
DB	474	ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGGAAGACGCCCGGTGG	533
QY	61	AspAspSerSerArgIleIleAsnGlySerAspCysAspMethIleThrGlnProTrp	80
DB	534	GATGACAGCAGCAGCGCATCATCATGATGATGATGATGATGATGATGATGATGATG	593
QY	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro	100
DB	594	CAGGCGCGCTGTGCTAGGCCACACAGCTCTACTGGGGCGGTGTGGTGATCCA	653
QY	101	GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis	120
DB	654	CAGTGGCTCTCAGCGCGCCACTGCAGGAAGAAAGTTTTCAGAGTCCGTCTCGGCCAC	713
QY	121	TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle	140
DB	714	TACTCCCTCTCACCAGTTTATGAATCTGGGCAGCAGATGTTCCAGGGGTCAAATCCATC	773
QY	141	ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn	160
DB	774	CCCCACCTTGGTACTCCACCCCTTACGACCTCTACGACCTATGCTCATCAACTGAAC	833
QY	161	ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer	180

DB	834	AGAAGAATTTCGTCCTCCACTAAAGATGTCAGACCCATCAACGTCCTCTCTCATTTCCCTCT	893
QY	181	AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe	200
DB	894	GCTGGGACAAAGT	953
QY	201	ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla	220
DB	954	CCTAAGGTCTCTCCAGTGTCTGAATATCAGCGTCTAAGTCAGAAAAAGGTGCGAGGATGCT	1013
QY	221	TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer	240
DB	1014	TACCCGAGACAGATAGATGACCATGTTCTGCGCGGTGACAAAAGCAGGTAGAGACTCC	1073
QY	241	CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer	260
DB	1074	TGCCAGGTGTATCTGCGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTC	1133
QY	261	TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys	280
DB	1134	TGGGGAGATTACCTTTGTCGCGGCCAACAGACCGGGTGTCTACACGAACCTTCTGCAAG	1193
QY	281	PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer	293
DB	1194	TTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC	1232

Search completed: August 5, 2003, 12:30:49

Job time : 282 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 5, 2003, 12:11:15 ; Search time 77 Seconds

(without alignments)  
1679.549 Million cell updates/sec

Title: US-09-936-271B-14

Perfect score: 1608

Sequence: 1 MATARPPWVWLICALITALL.....VTNLCCKTKWQIETIQANS 293

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xl  
-Q=/cgn2\_1/USPTO.spool/US09936271/runat\_04082003\_104246\_25727/app\_query.fasta\_1.455  
-DB=Issued Patents\_NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09936271@cgn\_1\_1\_48 -runat\_04082003\_104246\_25727 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1608	100.0	1570	4	US-09-996-243-308
2	1595	99.2	1476	2	Sequence 308, App
3	1595	99.2	1476	3	Sequence 2, Appli
4	1595	99.2	1476	4	Sequence 2, Appli
5	1403	87.3	1504	4	Sequence 1, Appli
6	725	45.1	735	4	Sequence 94, Appl
7	671	41.7	765	4	Sequence 524, App
8	654.5	40.7	1052	4	Sequence 10, Appl
9	652.5	40.6	1146	4	Sequence 247, App
10	652.5	40.6	1166	3	Sequence 7, Appli
11	650.5	40.5	833	2	Sequence 2, Appli
12	650.5	40.5	1192	3	Sequence 8, Appli

13	650.5	40.5	1292	4	US-09-205-258-189
14	650.5	40.5	1314	3	US-09-025-059-2
15	639.5	39.8	994	3	US-09-008-271A-19
16	638.5	39.7	944	3	US-09-070-526-1
17	638	39.7	871	1	US-08-744-026-2
18	638	39.7	871	2	US-09-102-732-2
19	638	39.7	871	3	US-09-261-767-2
20	638	39.7	871	4	US-08-969-987-7
21	638	39.7	1248	3	US-09-020-956-171
22	638	39.7	1248	3	US-09-030-607-171
23	638	39.7	1248	4	US-09-439-313-171
24	638	39.7	1248	4	US-09-352-616A-171
25	638	39.7	1248	4	US-09-232-149A-171
26	635	39.5	1037	4	US-09-386-642-60
27	632.5	39.3	1049	4	US-09-386-642-9
28	620	38.6	1214	4	US-09-439-313-225
29	620	38.6	1214	4	US-09-352-616A-225
30	620	38.6	1214	4	US-09-232-149A-225
31	620	38.6	1215	4	US-09-439-313-326
32	620	38.6	1215	4	US-09-352-616A-326
33	620	38.6	1215	4	US-09-232-149A-326
C 34	612.5	38.1	969	3	US-09-502-600-30
35	612.5	38.1	986	2	US-08-557-146-1
36	612.5	38.1	986	2	US-09-154-344-1
C 37	612.5	38.1	1089	3	US-08-930-188-1
C 38	612.5	38.1	1089	3	US-08-930-188-3
C 39	612.5	38.1	1089	5	PCT-US96-04294-1
C 40	612.5	38.1	1089	5	PCT-US96-04294-3
41	590	36.7	1386	2	US-08-897-340-1
42	590	36.7	1386	3	US-09-252-329-1
C 43	585	36.4	1364	4	US-09-280-116-20
44	582.5	36.2	1265	3	US-09-020-956-173
45	582.5	36.2	1265	3	US-09-030-607-173

ALIGNMENTS

RESULT 1  
US-09-996-243-308  
; Sequence 308, Application US/09996243  
; Patent No. 6478825  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C13  
; CURRENT APPLICATION NUMBER: US/09/996, 243  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16

PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16	PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17	PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22	PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-22	PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23	PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/090862

;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Alignment Scores:  
Pred. No.: 9,79e-161 Length: 1570  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-936-271B-14 (1-293) x US-09-996-243-308 (1-1570)

QY 1 MetAlaThrAlaArgProTirpMetTirpValLeuCysAlaLeuLeuLeuLeu 20  
DB 354 ATGGCTACAGCAAGCCCTCGATGGGTCTGTCTGTCTGTATCACAGCCTTGCCT 413  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
DB 414 CTGGGGTTCACAGACGATGTTCTGCCAACAAATGATGTTCTGTGACCACTCTTAAC 473  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
DB 474 ACCGTGCTCTGGGAGCAACAGGACCTGGGAGCTGGGCGGGGAGAGCCCGGTGCG 533  
QY 61 AspAspSerSerArgIleAlaHisCysArgGlySerAspCysAspMetHisThrGlnProTirp 80  
DB 534 GATCAGCAGCAGCCGATCATCAATGGATCCGATCGGATATGCACACCCCGGTGG 593  
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
DB 594 CAGGCGCGCTGTGCTTAAGGCCCAACAGCTCTACTCGGGGGGGTGTGGTGCATCCA 653  
QY 101 GlnTrpLeuThrAlaAlaHisCysArgGlySerValPheArgValArgLeuGlyHis 120  
DB 654 CAGTGGTGTCTCAGGGCGCCCTACTGAGGAAGAAAGTTTTCAGAGTCCGCTCGGCCAC 713  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
DB 714 TACTCCCTGTCCACAGTTTATGATCTGGGAGCAGATGTTCCAGGGGGTCAATCCATC 773  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAsn 160  
DB 774 CCCCACCTGGCTACTCCACCTGGCCACTTAAACGACCTCATCACTCAACTGAAC 833  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
DB 834 AGAAGAATTCCTCCCACTAAAGATGTCAGACCCCAACAGTCTCTCTCATTTGCCCTCT 893  
QY 181 AlaglyThrLysCysLeuValSerGlyTirpGlyThrLysSerProGlnValHisPhe 200  
DB 894 GCTGGGCAAAAGTCTGTGGTCTGGTGGGGGACAAAGAGCCCCCAAGTGCATTC 953  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220

DB 954 CTAAGGTCTCCAGTCTTGAATATCATCAGCGTGTCTAAGTACAGAAAGGTGGAGGATGCT 1013  
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
DB 1014 TACCCGAGACAGATAGATGACACCATGTTCTGCCCGGTGACAAAGCAGGTAGAGACTCC 1073  
QY 241 CysGlnGlyAspSerGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
DB 1074 TGCCAGGGTATTCTGGGGGCGCTGTGTCTGCAATGCTCCCTGCAGGACTCGTGTCC 1133  
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
DB 1134 TGGGAGATTACCTTGTGCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAG 1193  
QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
DB 1194 TTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC 1232

RESULT 2

US-08-824-874-2  
; Sequence 2, Application US/08824874  
; Patent No. 5962300  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,874  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1476 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: KERANOT02  
; CLONE: 820694  
US-08-824-874-2

Alignment Scores:  
Pred. No.: 2,12e-159 Length: 1476  
Score: 1595.00 Matches: 293  
Percent Similarity: 99.66% Conservative: 0  
Best Local Similarity: 99.66% Mismatches: 0  
Query Match: 99.19% Indels: 1  
DB: 2 Gaps: 0  
US-09-936-271B-14 (1-293) x US-08-824-874-2 (1-1476)

```
OY 1 MetaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuThrAlaLeuLeu 20
Db 277 ATGGCTACAGCAAGACCCCTGGATGGTGTCTGTCTGTATCATCAGCCCTTGGCTT 336
OY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 337 CTGGGGTCTACAGAGCATGTTCTCGCAACAATGATGTTTCTGTGACCACTTAAAC 396
OY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 397 ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGAGAGCGCCGCTG 456
OY 61 AspAspSerSerArgIleAAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 457 GATGACAGCAGCAGCCGATCATCAATGGATCGACTCGATATGCACACCCAGCCGTGG 516
OY 81 -GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPr 100
Db 517 GCAGGCGCGCTGTGTCTAAGGCCCCACCACTCTACTGCGGGCGGTGTGGTGCATCC 576
OY 100 oGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyH 120
Db 577 ACAGTGGCTGCTACGGCGCCGCTCGAGAGAAAGTTTTCAGAGTCCGCTCGGCCA 636
OY 120 sTyrSerLeuSerProValTyrGluSerGlyGlnMetPheGlnGlyValLysSerIl 140
Db 637 CTACTCCCTGTCCACAGTTTATGAATCTGGGAGCAGATGTTCCAGGGGTCAAAATCCAT 696
OY 140 eProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAs 160
Db 697 CCCCCACCTGGCTACTCCACCTGGCCACTCTAAGCAGCTCTAAGTCAAGAACTGAA 756
OY 160 nArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProse 180
Db 757 CAGAGAAATTCCTCCACTAAGATGTCAGACCATCAACGTCTCTCTATTGTCCTC 816
OY 180 rAlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPh 200
Db 817 TGCTGGGACAAAGTGTGTGTGCTGGGCTGGGGGACCAACCAAGAGCCCAAGTGCATT 876
OY 200 eProLysValLeuGlnCysLeuAsnIleSerValLeuSerClnLysArgCysGluAspAl 220
Db 877 CCCTAAGGTCTCCAGTGTCTGAATATACAGCGTGTCTAAGTCAAGAAAGGTCGAGATGC 936
OY 220 aTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspse 240
Db 937 TTACCCGAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGACTC 996
OY 240 rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValse 260
Db 997 CTGCCAGGGTGATTCCTGGGGGCGCTGTGTCTGCAATGGCTCCCTGCGAGGACTCGTGC 1056
OY 260 rTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLy 280
Db 1057 CTGGGGAGATTACCTCTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAA 1116
OY 280 sPheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 1117 GTTCACCAAGTGGATCCAGAAACCATCCAGGCCCAACTCC 1156
```

## RESULT 3

```
US-09-210-084-2
; Sequence 2, Application US/09210084
; Patent No. 6197511
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
```

```
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,084
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1476 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: KERANOT02
; CLONE: 820694
```

US-09-210-084-2

```
Alignment Scores:
Pred. No.: 2,12e-159 Length: 1476
Score: 1595.00 Matches: 293
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 0
Query Match: 99.19% Indels: 1
DB: 3 Gaps: 0
```

US-09-936-271b-14 (1-293) x US-09-210-084-2 (1-1476)

```
OY 1 MetaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuThrAlaLeuLeu 20
Db 277 ATGGCTACAGCAAGACCCCTGGATGGTGTCTGTCTGTATCATCAGCCCTTGGCTT 336
OY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 337 CTGGGGTCTACAGAGCATGTTCTCGCAACAATGATGTTTCTGTGACCACTTAAAC 396
OY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 397 ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGAGAGCGCCGCTG 456
OY 61 AspAspSerSerArgIleAAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 457 GATGACAGCAGCAGCCGATCATCAATGGATCGACTCGATATGCACACCCAGCCGTGG 516
OY 81 -GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPr 100
Db 517 GCAGGCGCGCTGTGTCTAAGGCCCCACCACTCTACTGCGGGCGGTGTGGTGCATCC 576
OY 100 oGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyH 120
Db 577 ACAGTGGCTGCTACGGCGCCGCTCGAGAGAAAGTTTTCAGAGTCCGCTCGGCCA 636
OY 120 sTyrSerLeuSerProValTyrGluSerGlyGlnMetPheGlnGlyValLysSerIl 140
Db 637 CTACTCCCTGTCCACAGTTTATGAATCTGGGAGCAGATGTTCCAGGGGTCAAAATCCAT 696
OY 140 eProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAs 160
```

697	Db		CCCCCACCCTGGGTACTCCACCCCTGGCCACTTAACGACCTCATGCTCATCAAACTGAA	756
160	QY		nArqArqIleArqProThrLysAspValArgProIleAsnValSerSerHisCysProse	180
757	Db		CAGAAGAAATCGTCCCACTAAAGATGTCAGACCCATCAAGCTCTCTCTATTGTGCTC	816
180	QY		rAlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPh	200
817	Db		TGCTGGGACAAAGTGCTTGCTGTCGCGCTGGGGGCAACCAAGAGCCCCCAAGTGCAC	876
200	QY		eProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAl	220
877	Db		CCCTAAGGTCCTCCAGTGTCTGAATATCAGCGTCTAAGTCAGAAAAAGGTGCGAGG	936
220	QY		aTyrProArqGlnIleAspThrMetPheCysAlaGlyAspLysAlaGlyArgaspSe	240
937	Db		TTACCCGAGACACATAGATGACACCATGTCTCGCGCGGTGACAAAGCAGGTAGAC	996
240	QY		rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValse	260
997	Db		CTGCCAGGTGATTCTGGGGGGCTGTGGTCTGCATGGCTCCTCGAGGGACTCTGTG	1056
260	QY		rTrpGlyAspTyrProCysAlaAlaArgProAsnArgProGlyValTyrThrAsnLeuCy	280
1057	Db		CTGGGAGATTACCTTGTCGCGCGGCCAACACACCGGGTGCTACACGAACCTCTG	1116
280	QY		sPheThrLysTyrPileGlnGluThrIleGlnAlaAsnSer	293
1117	Db		GTTCCACCAAGTGGATCCAGAAACCATCCAGGCCAACTCC	1156

## RESULT 4

US-09-764-762-2

; Sequence 2, Application US/09764762

; Patent No. 6472195

**GENERAL INFORMATION:**

APPLICANT: Hillman, Jennifer L.

.....

1. TITLE OF INVENTION: NOVEL KALLIKREIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte pharmaceuticals, Inc.

ADDRESSEE: INCYTE PHARMAC  
STREET: 3174 Porter Drive

SIREET: 31/4 FO  
CITY: palo Alto

CITY: PALO  
STATE: CASTATE: CA  
COUNTRY: USA

COUNTRY: U  
ZTP: 94304

ZIP: 94304  
COMPUTER READING FORM.

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

```

;
OPERATING SYSTEM: DOS

```

```

;
SOFTWARE: FastSEQ for Windows Ver

```

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/00

FILING DATE: 16-Jan-2001

; CLASSIFICATION: &lt;U&gt;

;  
;  
PRIOR APPLICATION DATA:

APPLICATION NUM:

## ; FILING DATE: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REGISTRATION NUMBER: 3077  
REFERENCE/DOCKET NUMBER:

: REFERENCE/DOCKET NUMBER  
: TELECOMMUNICATION INFORMATION

TELEPHONE: 415-855-0559

TELEPHONE: 415-855-0333  
TELEFAX: 415-845-4166

TELEFAX: 413-843-418  
: INFORMATION FOR SEO ID NO: 3:

; INFORMATION FOR SEQ ID NO: 2:  
: SEQUENCE CHARACTERISTICS:  
:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1476

LENGTH: 14/6 base pair

TYPE: nucleic acid

STRADEDNESS

```

; LIBRARY: KERANOT02
; CLONE: 820694
; SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-764-762-2

Alignment Scores:
Pred. No.: 2 12e-159      Length: 1476
Score: 1595.00      Matches: 293
Percent Similarity: 99.66%      Conservative: 0
Best Local Similarity: 99.66%      Mismatches: 0
Query Match: 99.19%      Indels: 1
DB: 4      Gaps: 0

```

US-09-936-271B-14 (1-293) x US-09-764-762-2 (1-1476)

Qy	1	MetaLThrAlaArgProProTrrPmetTrrpValLeuCysAlaLeuIleThrAlaLeuLeu	20
Db	277	ATGGCTACACGAAGACCCCGCTGGATGTGGGTGCTGTGCTCTGATCAGACGCTTGCTT	336
Qy	21	LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn	40
Db	337	CTGGGGTCACAGAGATGTTCTGCCCAACAATGATGTTTCTGTGACCACTCTTAAC	396
Qy	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer	60
Db	397	ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCCGGGAAGACGCCGGTCG	456
Qy	61	AspAspSerSerArgIleIleLeuAsnGlySerAspCysAspMetHisThrGlnProTrrp	80
Db	457	GATCAGCAGCAGCGCGCATCATCAATGGATCGGACTGCGATATGCACACCCAGCGGTG	516
Qy	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPr	100
Db	517	GCAGGCGCGCTGTGTAAAGGCCAACCAACAGCTCTACTGCGGGCGGTGTGGTGATCC	576
Qy	100	oGlnTrrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHi	120
Db	577	ACAGTGGCTCTCACGGCGCCCACTGCAGAGAAGAAGTTTTCAGAGTCTGCTCGGCCA	636
Qy	120	sTrrSerLeuSerProValTrrGluSerGlyGlnGlnMetPheGlnGlyValLysSerII	140
Db	637	CTACTCCCTCTCACCAAGTTTATGAATCTGGGCAGCAGATGTTCCAGGGGTCAATCCAT	696
Qy	140	eProHisProGlyTrrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAs	160
Db	697	CCCCCACCTGGCTACTCCCACTGGCCACTCTAACGACCTCATGCTCATCAACATGAA	756
Qy	160	nArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProse	180
Db	757	CAGAAGAATTCGTCCCACTAAAGATGTCACAGCCATCAACGTCTCTCTCATTTGCCCTC	816
Qy	180	rAlaGlyThrLysCysLeuValSerGlyTrrpGlyThrThrLysSerProGlnValHisPh	200
Db	817	TGCTGGGACAAAGTGCTGTGGTCCGGCTGGGGGACAACCAAGAGCCCCCAAGTGACATT	876
Qy	200	eProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAl	220
Db	877	CCCTAAGGTCTCCAGTGTGTAATATCAGGTGCTTAAGTCAGAAAAGGTGCCAGGATGC	936
Qy	220	aTrrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSe	240
Db	937	TTACCCGAGACATAGATGACACCATGTTCTGCGCGGTGCACAAAGCAGGTAGAGACTC	996
Qy	240	rCysGlnGlyAspSerGlyProValValCysAsnGlySerLeuGlnGlyLeuValSe	260
Db	997	CTGCCAGGGTGATCTCTGGGGGGCTGTGGTGTCTGCAATGGCTCCCTGCGAGGGACTCTG	1056
Qy	260	rTrrpGlyAspTrrProCysAlaArgProAsnArgProGlyValTrrThrAsnLeuCysLy	280
Db	1057	CTGGGAGATTACCTTGTCCCGGGCCCAACAGCCGGGTGCTTACACGAACCTCTGC	1116
Qy	280	sPheThrLysTrrpIleGlnLuthrIleGlnAlaAsnSer	293

Db 1117 GTTCAACCAAGTGGATCCAGGAAACCATCCAGGCCCAACTCC 1156

## RESULT 5

US-09-280-116-1  
; Sequence 1, Application US/09280116A  
; Patent No. 6331427  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs  
; FILE REFERENCE: 5800-24, 035800/176965  
; CURRENT APPLICATION NUMBER: US/09/280.116A  
; CURRENT FILING DATE: 1999-03-26  
; NUMBER OF SEQ ID NOS: 268  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1504  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: trypsin-like serine proteases  
US-09-280-116-1

Alignment Scores:  
Pred. No.: 4, 79e-139 Length: 1504  
Score: 1403.00 Matches: 280  
Percent Similarity: 93.77% Conservative: 6  
Best Local Similarity: 91.80% Mismatches: 7  
Query Match: 87.25% Indels: 12  
DB: 4 Gaps: 2

US-09-936-271B-14 (1-293) x US-09-280-116-1 (1-1504)

Qy 1 MetAlaThrAlaArgProTropMetTropValLeuCysAlaLeuIleThrAlaLeu 20  
Db 277 ATGGCTACAGCAAGACCCCTGGTGTGGTGTCTGTCTGTGATCAGACGCTTGGT 336  
Qy 21 LeuGlyValThrGluHisValLeuAla-AsnAspValSerCysAspHisProSerAs 40  
Db 337 CTGGGGGTACAGAGCATGTTTTCGCCAAACAATGATGTTTCTGTGACCACTCTAA 396  
Qy 40 nThr-ValProSerGlySer-AsnGlnAspLeuGlyAla---GlyAlaGlyGluAspAla 58  
Db 397 CACCGTGGCCCTCTGGGAGCAACACGAGCTTGGGGAACTGGGGCCCGGGGAAGACGCC 456  
Qy 59 ArgSerAspSerSerSer-ArgIleIleAsnGlySerAsp-CysAspMetHisThrG 78  
Db 457 CGGTGGATGACAGCAGCCCGCATCATCAATGATCCGAACTCGGATATGCACACC 516  
Qy 78 InProTropGlnAla---AlaLeuLeuLeuArgProAsnGlnLeu--TyrCysGlyAlaVa 96  
Db 517 AGCCGTGGCAGCCCGCGCTGTTTAAAGCCGCCCAACCAAGTCTTATTGCGGGCGGT 576  
Qy 96 LeuValHisProGlnTrpLeuLeuThrAlaAla-HisCysArgLysLysValPheArgV 116  
Db 577 GTTGGTCATCCACAGTGGTGTTCAGCGGCCCGCCCTACTGCAGAGAAAGTTTTCAG 636  
Qy 116 alaArgLeuGlyHisTyrSerLeuSerProValTyrGluSerGlyGlnGln-MetPheGln 135  
Db 637 TCCGTCTCGGCCACTATTTCCTGTGCACAGTTTATGAATCTGGCAGCAAGATGTCAG 696  
Qy 136 GlyValLysSerIleProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMet 155  
Db 697 GGGGTCAATCATCCCGCCCGCTGGTACCTCCACCCCTGGCCACTTAAGGACCTCATG 756  
Qy 156 LeuIleLysLeuAsnArgArgIleArgProThrLysAspValArgProIleAsnValSer 175  
Db 757 CTCATCAACTGACAGAGAGATTCGTCCTCACTAAAGATGTCAGACCCATCAAGCTTCC 816  
Qy 176 SerHisCysProSerAlaGlyThrLysCysLeuValSerGlyTropGlyThrThrLysSer 195  
Db 817 TCTCATTTCTCCTCTGCTGGGACAAAGTCTGTGGTGTCTGCTGGGGGACAAACAGAGC 876  
Qy 196 ProGlnValHisPheProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLys 215

Db 877 CCCCAAGTGCACCTTCCCTAAGGTCTCCAGTGTGTAATATACAGCTGCTAAGTCAGAAA 936  
Qy 216 ArgCysGluAspAlaTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLys 235  
Db 937 AGGTGCGAGGATGCTTACCGAGACACATAGATGACACCATGTTCTCGCGCGGTGACAAA 996  
Qy 236 AlaGlyArgAspSerCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeu 255  
Db 997 GCAGGTAGAGACTCCTGCCAGGGTGATTCCTGGGGGGCTGTGTGCTGCAATGGCTCCCTG 1056  
Qy 256 GlnGlyLeuValSerTyrGlyAspTyrProCysAlaArgProAsnArgProGlyValty 275  
Db 1057 CAGGACTGTGTCTCGGGGAGATTACCTTGTGCCCGGCCCAACAGACCGGTGTCTA 1116  
Qy 275 rThrAsnLeuCysLysPheThrLysTrpIleGlnGlnThrIleGlnAlaAsnSer 293  
Db 1117 CACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCCAACTCC 1171

## RESULT 6

US-09-602-877A-94/c  
; Sequence 94, Application US/09602877A  
; Patent No. 6432707  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.446C5  
; CURRENT APPLICATION NUMBER: US/09/602.877A  
; CURRENT FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 94  
; LENGTH: 735  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-602-877A-94

Alignment Scores:  
Pred. No.: 1, 09e-67 Length: 735  
Score: 725.00 Matches: 130  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 45.09% Indels: 0  
DB: 4 Gaps: 0

US-09-936-271B-14 (1-293) x US-09-602-877A-94 (1-735)

Qy 164 ArgProThrLysAspValArgProIleAsnValSerSerHisCysProSerAlaGlyThr 183  
Db 734 CGTCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCATTTCTCTCTGCTGGGACA 675  
Qy 184 LysCysLeuValSerGlyTropGlyThrThrLysSerProGlnValHisPheProLysVal 203  
Db 674 AAGTGTCTGTGTCTGTGGGACCAACCAAGAGCCCAAGTGTCTCTTCCCTTAAGGTC 615  
Qy 204 LeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAlaTyrProArg 223  
Db 614 CTCAGTGTCTGAATATACAGCTGTAAAGTCAGAAAGGTGCGAGGATGCTTACCCGAGA 555  
Qy 224 GlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSerCysGlnGly 243  
Db 554 CAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTCTGCCAGGT 495  
Qy 244 AspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSerTropGlyAsp 263  
Db 494 GATTCTGGGGGCGCTGTGGTGTGCAATGGCTCTCCCTGCGAGGACTCGTGTCTCTGGGAGAT 435  
Qy 264 TyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLysPheThrLys 283  
Db 434 TACCCTTGTGCCCGGCCCAACAGACACCGGGTGTCTACAGAACCTCTGCAAGTTCCACCAAG 375

QY 284 TptleGlnGluThrIleGlnAlaAsnSer 293  
 |||||  
 Db 374 TGGATCCAGGAACCATCCAGGCCAACTCC 345

## RESULT 7

US-09-439-313-524  
 ; Sequence 524, Application US/09439313  
 ; Patent No. 6329505  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan Louise  
 ; APPLICANT: Jiang Yuqi  
 ; APPLICANT: Reed, Steven G.  
 ; APPLICANT: Kalos, Michael  
 ; APPLICANT: Fanger, Gary  
 ; APPLICANT: Retter, Mark  
 ; APPLICANT: Solk, John  
 ; APPLICANT: Day, Craig  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
 ; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
 ; FILE REFERENCE: 210121.427C9  
 ; CURRENT APPLICATION NUMBER: US/09/439, 313  
 ; CURRENT FILING DATE: 1999-11-12  
 ; NUMBER OF SEQ ID NOS: 575  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 524  
 ; LENGTH: 765  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 US-09-439-313-524

Alignment Scores:  
 Pred. No.: 6,11e-62 Length: 765  
 Score: 671.00 Matches: 134  
 Percent Similarity: 61.77% Conservative: 47  
 Best Local Similarity: 45.73% Mismatches: 72  
 Query Match: 41.73% Indels: 40  
 DB: 4 Gaps: 6

US-09-936-271B-14 (1-293) x US-09-439-313-524 (1-765)

QY 1 MetAlaThrAlaArgProTtpMetTtpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
 |||||  
 Db 1 ATGCCACAGCAAGAAATCCCTGGGCTGGTTCTCTGGGTACCTCATC----- 48  
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
 |||||  
 Db 49 CTGGTGTCTC-----GCAGGA 63  
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
 |||||  
 Db 64 TCGCTGCTCTCTGCT----- 78  
 QY 61 AspAspSerSerSerArgIlelleAsnGlySerAspCysAspMethIsthrGlnProTtp 80  
 |||||  
 Db 79 -----AGTGCAGCCAAATCATAAACGGCGAGGACTGCAGCCCGCATCGACCGCTGG 132  
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
 |||||  
 Db 133 CAGCGCGCACTGCTATGNA-----AACGAATTTCTCTCGGGCGTCTGGTGATCCG 199  
 QY 101 GlnTtpLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLysHis 120  
 |||||  
 Db 190 CAGTGGTGTCTGCAGCCGCACACTGTTCCAGAACTCTACACCATCGGCTGGGCGCTG 249  
 QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnMetPheGlnGlyValLysSerIle 140  
 |||||  
 Db 250 CACAGTCTTGAGGCCCAACCAAGAGCCAGGAGCAGATGTGTGGAGCCAGCCTCTCCGTA 309  
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160

Db 310 CGGCACCCAGAGTACACACACCCCTGCTCGCTAACGACCTCATGCTCATCAAGTTGGAC 369  
 |||||  
 QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
 |||||  
 Db 370 GAATCGGTGTCGAGTCTGACACCATCCGGAGCATCAGCATTTGCTCGCAGTGCCTTACC 429  
 |||||  
 QY 181 AlaGlyThrLysCysLeuValSerGlyTtpGlyThrThrLysSerProGlnValHisPhe 200  
 |||||  
 Db 430 GCGGGGAACCTCTTGGCTCGCTTCTGCTGGGGTCTGCTGGCAACGCGCAAGTG----- 483  
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
 |||||  
 Db 484 CTTACCGTGTGTCAGTCCGTGAACGCTCGTGGTGTCTGAGGAGGTCTCGCAGTAAGCTC 543  
 |||||  
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAla---GlyAspLysAlaGlyArgAsp 239  
 |||||  
 Db 544 TATGACCCGCTGTATCACCCACCCAGCATGTTCTGGCGCGCGGAGGCAAGACCAAGAGGAC 603  
 |||||  
 QY 240 SerCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuVal 259  
 |||||  
 Db 604 TCTGCAACGGTGACTCTGGGGGGCCCTCATCTGCAACGGGTACTTGCAGGGCTTGTG 663  
 |||||  
 QY 260 SerTtpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys 279  
 |||||  
 Db 664 TCTTTCGAAAGACCCCGTGTGCCAAGTTGGCGTGCAGGTGTCTACACCAACCTCTGC 723  
 |||||

## RESULT 8

US-09-386-642-10  
 ; Sequence 10, Application US/09386642  
 ; Patent No. 6420157  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Darrow, Andrew  
 ; APPLICANT: Qi, Jensen  
 ; APPLICANT: Andrade-Gordon, Patricia  
 ; TITLE OF INVENTION: Zymogen Activation System  
 ; FILE REFERENCE: ORT-1028  
 ; CURRENT APPLICATION NUMBER: US/09/386, 642  
 ; CURRENT FILING DATE: 1999-08-31  
 ; NUMBER OF SEQ ID NOS: 60  
 ; SOFTWARE: PatentIn ver. 2.0  
 ; SEQ ID NO 10  
 ; LENGTH: 1052  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene  
 ; OTHER INFORMATION: with homo sapien serine protease catalytic domain  
 US-09-386-642-10

Alignment Scores:  
 Pred. No.: 5,54e-60 Length: 1052  
 Score: 654.50 Matches: 129  
 Percent Similarity: 62.99% Conservative: 48  
 Best Local Similarity: 45.91% Mismatches: 87  
 Query Match: 40.70% Indels: 17  
 DB: 4 Gaps: 7

US-09-936-271B-14 (1-293) x US-09-386-642-10 (1-1052)

QY 19 LeuLeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisPro 38  
 |||||  
 Db 49 CTCTGCTGCTGGTGTCAATCTACTCTTGTGCCAGGTGTGGTCTCGACTACAAG 108  
 QY 39 SerAsnThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAla 58  
 |||||  
 Db 109 -----GACGAGCAGCAGCTGGAGCGGCGCT-----CTTGCT 141  
 QY 59 ArgSerAspAspSerSerArgIlelleAsnGlySerAspCys---AspMethIsthr 77







```
Query Match: 40.58% Indels: 7
DB: 3 Gaps: 4
US-09-936-271B-14 (1-293) x US-08-944-483-7 (1-1166)
Qy 65 SerArgIleLeuAsnGlySerAspCysAspMetHisThrGlnProTrpGlnAlaLeu 84
Db 160 ACCAGGATCATCAAGAGGTTCGAGTGCNAGCCCTCACTCCAGGCCCTGGCAGGAGCCCTG 219
Qy 85 LeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisProGlnTrpLeu 104
Db 220 TTCRAGAAAR---ACCGCGCTACTCTGTGGGGGACGCTCATCGCCCGCATGGCTCTG 276
Qy 105 ThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHisTyrSerLeu 124
Db 277 ACAGCAGCCCACTCCCTCAAGCCCGCTACATAGTTCACCTGGGGCAGCACAACCTCCAG 336
Qy 125 ProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerIleProHisPro 144
Db 337 AAG---GAGGAGGCTGTGAGCAGACCCGCGACCCACTGAGTCTTCCCGCCCGCCG 393
Qy 145 TyrSer-----HisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
Db 394 TTCACACACAGCCTCCCAACAAAGACCCGCAATGACATCATCTGTGTGAAGATGCA 453
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysPro 180
Db 454 TCGCCAGTCTCCATCACTGCGGTGTGCGACCCCTCACCTCTCTCTCACGCTGTGTCACT 513
Qy 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200
Db 514 GCTGGCAGCAGTCCCTCATTTTCGGGTGGGGCAGCAGTCCAGCCCGCCAGTTCAGCCTG 573
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGlyAspAla 220
Db 574 CCTCACACCTTGGCATCGCCCAACATCACCATCATTCAGCACCAGAGTGTGAGAAGCC 633
Qy 221 TyrProArgGlnIleAspThrMetPheCysAlaGly---AspLysAlaGlyArgAsp 239
Db 634 TACCCCGGCACATCACAGACACCATGTGTGTGCGCAGCGTGTGCGAGGAGGGGCAAGGAC 693
Qy 240 SerCysGlnGlyAspSerGlyProValValCysAsnGlySerLeuGlnGlyLeuVal 259
Db 694 TCCTGCCAGGTGACTCGGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 753
Qy 260 SerTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeu 279
Db 754 TCCTGGGCGCAGGATCCGTGTGCGATCACCGGAAAGCGTGTGTGTGTGTGTGTGTGTGT 813
Qy 280 LysPheThrLysTrpIleGlnGlnThrIleGlnAlaAsn 292
Db 814 AAATATGTGGACTGGATCCAGGAGACGATGAAGAACAAT 852
RESULT 11
US-08-790-137-2
; Sequence 2, Application US/08790137
; Patent No. 5840871
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL PROSTATE-ASSOCIATED
; TITLE OF INVENTION: KALLIKREIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/790,137
FILING DATE: Filed Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0195 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-790-137-2
Alignment Scores:
Pred. No.: 1,03e-59 Length: 833
Score: 650.50 Matches: 127
Percent Similarity: 58.62% Conservative: 43
Best Local Similarity: 43.79% Mismatches: 77
Query Match: 40.45% Indels: 43
DB: 2 Gaps: 7
US-09-936-271B-14 (1-293) x US-08-790-137-2 (1-833)
Qy 8 TrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeuGlyValThrGluHisVal 27
Db 30 TGGTTCCTGTTCTGTGTC-----CTCGCCCTGTCCCTGGGG----- 65
Qy 28 LeuAlaAsnAsnAspValSerCysAspHisProSerAsnThrValProSerGlySerAsn 47
Db 65 ----- 65
Qy 48 GlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSerAspSerSerArgIle 67
Db 66 -----GGGACTGTGTGTGCCCCCGGATTCAG-----TCCCGGATT 101
Qy 68 IleAsnGlySerAspCysAspMetHisThrGlnProTrpGlnAlaLeuLeuArg 87
Db 102 GTGGGAGGCTGGGAGTGTGAGCAGCATTCGCCAGCCCTGGCAGCGGCT---CTGTACCAG 158
Qy 88 ProAsnGlnLeuTyrCysGlyAlaValLeuValHisProGlnTrpLeuLeuThrAlaAla 107
Db 159 AAGACGGCGGTACTCTCTGTGGGGGACGNTCATNGCCCGCCAGATGTTCTCTGACAGACCC 218
Qy 108 HisCysArgLysLysValPheArgValArgLeuGlyHisTyrSerLeuSerProValTyr 127
Db 219 CACTGCTNAGCCCGCTCATATGATTCACCTGGGGCAGCACACCTCCAGAAAG---GAG 275
Qy 128 GluSerGlyGlnGlnMetPheGlnGlyValLysSerIleProHisProGlyTyrSer--- 146
Db 276 GAGGGCTGTGAGCAGACCCGCGCAGCAGCTGAGTCTTCCCGCCCGCGGCTTCAACAAC 335
Qy 147 -----HisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsnArgIle 163
Db 336 AGCCTCCCAACAAAGACCCCGCAATGACATCATCTGTGTGTGTGTGTGTGTGTGTGTGT 395
Qy 164 ArgProThrLysAspValArgProIleAsnValSerSerHisCysProSerAlaGlyThr 183
Db 396 TCCATCACCTGGGCTGTGCGACCCCTCACCTCTCTCTCACCTGTGTGTGTGTGTGTGTGT 455
Qy 184 LysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPheProLysVal 203
Db 456 AGCTGCCTCATTTCCGGCTGGGGCAGCAGCTCCAGCCCGCCCGCTTACGCTTACGCTC 515
```



[illegible]

```

; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 189
; LENGTH: 1292

```

Alignment Scores:		
Pred. No.:	2, 01e-59	Length:
Score:	850.50	Matches:
Percent Similarity:	68.24%	Conservative:
Best Local Similarity:	50.64%	Mismatches:
Query Match:	40.45%	Indels:
DB:	4	Gaps:

US-09-936-271B-14 (1-293) x US-09-205-258-189 (1-1292)

Qy	65	SerArglleleleAenGlySerAspCysAspMethHisThrGlnProTrpGlnAlaIateu	84
Db	264	ACCGAGTATCAAGAGGGTTCTGAGTGAAGGCTCTACTCCAGCCCTGCGAGCAGCC---	320
Qy	85	LeuLeuArgProAsnGlnLeuTyTCysGlyAlaValLeuValHisProGlnTrpLeu	104
Db	321	CTGTTCCAGAAAGACCGCGGTACTCTGTGGGGCGAGGCTCATCGCCCCAGATGGCTCTCG	380
Qy	105	ThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHisTyrSerLeuSer	124
Db	381	ACAGCAGCCCACTGCCTCAAGCCCGCTTACATAGTTCACTTGGGGCAGCACAACTCCAG	440
Qy	125	ProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerIleProHisProGly	144
Db	441	AAG---GAGGAGGGCTGTGAGCAGCAGCCGGACCGGACCGCTAGTCTCTCCCCACCCCGGC	497
Qy	145	TyrSer-----HisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn	160
Db	498	TTCAACAACAGCCTCCCCCAACAAGACCCAGCAATGACATCATCTGGTGAAGATGGCA	557
Qy	161	ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer	180
Db	558	TGCGCAGTCTCCATCAGCTGGCGTGTGGAGCCCTCACCCCTCTCTCTCAGCTGTGTCACT	617
Qy	181	AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe	200
Db	618	GCTGGCACCAAGCTGCTCATTTCCGGCTGGGGCAGCAGCTCCAGCCCCAGTTACGCGCT	677
Qy	201	ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspIa	220
Db	678	CCTACACCTTCGATCGCCAGCAATCAACATCACTAGCACCAGCAAGTGTGAGAACGCC	737
Qy	221	TyrProArgGlnIleAspAspThrMetPheCysAlaGly---AspLysAlaGlyArgAsp	239
Db	738	TACCCCGGCACATCACAGACACCATGGTGTGTGCCCGCTGCAGGAAGGGGCGCAAGGAC	797
Qy	240	SerCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuVal	259
Db	798	TCCTGCCAGGGTGACTCCGGGGGGCCCTCTGGTCTGTAAACCAGTCTCTTCAAGGCATATT	857
Qy	260	SerTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys	279
Db	858	TCCTTGGGGCCAGGATCCGTGTGGCATCAACCCGAAGGCTGTGTCTACACGAAAGTCTGCG	917
Qy	280	LysPheThrLysTrpIleGlnGluThrIleGlnAlaAsn	292
Db	918	AAATATGTGGACTCGATCCAGCAGACGATGAAGAACAAAT	956

RESULT 14  
US-09-025-059-2  
; Sequence 2, Application US/09025059  
; Patent No. 6075136



```
;
;
; IMMEDIATE SOURCE:
;
; LIBRARY: COLNNOT27
; CLONE: 1798496
; SEQUENCE DESCRIPTION: SEQ ID NO: 19 :
US-09-008-271A-19
```

```
Alignment Scores:
Pred. No.: 1.97e-58 Length: 994
Score: 639.50 Matches: 126
Percent Similarity: 57.09% Conservative: 43
Best Local Similarity: 42.57% Mismatches: 80
Query Match: 39.77% Indels: 47
DB: 3 Gaps: 7
```

US-09-936-271B-14 (1-293) x US-09-008-271A-19 (1-994)

```
Qy 1 MetAlaThrAlaArgPro-----ProTrpMetTrpValLeuCysAlaLeuIle 16
Db 155 ATGGGACGCGCCGACCTCGTGGCGCAAGACGCTGGATGTTCTCG-----199
Qy 17 ThrAlaLeuLeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAsp 36
Db 200 -----CTCTGTGGTGGG-----211
Qy 37 HisProSerAsnThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGlu 56
Db 212 -----GGAGCCTGGCGAGGACAC 229
Qy 57 AspAlaArgSerAspAspSerSerArgIleAsnGlySerAspCysAspMethHis 76
Db 230 TCCAGGCGACAGAGGAGAC-----AAGTGTGGGGGTGATGATGCCAACCCCAT 280
Qy 77 ThrGlnProTrpGlnAlaLeuLeuLeuArgProAsnGlnLeuTyrcysGlyAlaVal 96
Db 281 TCGCAGCCTTGGCAGGCGGCC---TTGTCCAGGCGCAGCACTCTCTGTGGCGGTGC 337
Qy 97 LeuValHisProGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgVal 116
Db 338 CTTGTAGTGGCACTGGTCCITTACAGCTGCCACTGTAAACCCGAAATACACAGTA 397
Qy 117 ArgLeuGlyHisTySerLeuSerProValTyroGluSerGlyGlnGlnMetPheGlnGly 136
Db 398 CGCCTGGGAGACACACAGCCATACAGAAAT---AAAGATGGCCCGAGCAAGAAATACCTGTG 454
Qy 137 ValLysSerIleProHisProGlyTySerHis-----ProGlyHisSerAsnAsp 153
Db 455 GTTCAGTCCATCCACACCCCTGCTACACAGCAGCGATGTGGAGGACCAACACCATGAT 514
Qy 154 LeuMetLeuIleLysLeuAsnArgArgIleArgProThrLysAspValArgProIleAsn 173
Db 515 CTGATGCTTCTCAACTGGGTGACCGAGCATCCCTGGGGTCCAAAGTGAAGCCCATCAGC 574
Qy 174 ValSerSerHisCysProSerAlaGlyThrLysCysLeuValSerGlyTrpGlyThrThr 193
Db 575 CTGGCAGATCATTCACCCAGCCCTGGCGCAGAAAGTGCACCGCTCTCAGCTGGGGCAGTGC 634
Qy 194 LysSerProGlnValHisPheProLysValLeuGlnCysLeuAsnIleSerValLeuSer 213
Db 635 ACCAGTCCCGGAGAGAAATTTCTCGACACTCTCAACTGTGCGAAGTAAATACTTTCCC 694
Qy 214 GlnLysArgCysGluAspAlaTyProArgGlnIleAspThrMetPheCysAlaGly 233
Db 695 CAGAAGAAGTGTGAGGATGCTTACCGGGCGAGATCACAGATGCATGGTCTGTGCAGGC 754
Qy 234 AspLysAlaGlyArgAspSerCysGlnGlyAspSerGlyGlyProValValCysAsnGly 253
Db 755 AGCAGCAAAAGGGGTGACACACCTGCGAGCGGATTTCTGGAGGCCCTCGGTGTGTGATGGT 814
Qy 254 SerLeuGlnGlyLeuValSerTrpGlyAspTyProCysAlaArgProAsnArgProGly 273
Db 815 GCACCTCCAGGGCATCACATCTCTGGGGCTCAGACCCCTGTGGGAGGTTCGGACAAACCTGGC 874
```

```
Qy 274 ValTyThrAsnLeuCysLysPheThrLysTrpIleGlnGluThrIle 289
Db 875 GTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATCATA 922
```

Search completed: August 5, 2003, 14:00:12  
Job time : 84 secs